

FORMACIÓN PROFESIONAL · INVESTIGACIÓN COMERCIAL · UNIDAD 11

Base de datos relacional

Lunes 31 de mayo de 2027. Las encuestas del Estudio de kits 2027 están en hojas de cálculo sueltas y Unai propone una hoja grande, una fila por respuesta, con el nombre, la dirección y el teléfono de quien contesta. ¿Qué representa cada fila y cuántas personas han contestado de verdad?

PARA EMPEZAR

En esta unidad aprenderás a...



Identificar entidades, tablas, registros, campos, claves y relaciones, y saber qué representa cada fila.



Diseñar una base de datos relacional con claves primarias y foráneas, y comprobar su integridad.



Ejecutar consultas de selección, agregación, unión y acción, y contar personas, encuestas y respuestas sin confundirlas.



Construir formularios con validación y probarlos con datos válidos e inválidos.



Proteger, copiar, restaurar y mantener la base, y crear informes, etiquetas y exportaciones.

SITUACIÓN DE PARTIDA

«Una hoja grande, una fila por respuesta»

Lunes 31 de mayo de 2027. Las encuestas del **Estudio de kits 2027** de **La Despensa de Nora** (caso ficticio) están en hojas de cálculo sueltas, y Nora quiere una sola base para el informe de junio. Unai lo tiene claro: «Una hoja grande, una fila por respuesta, y en cada fila el nombre, la dirección y el teléfono de quien contesta».

Adrián, que pasa a limpio los cuestionarios, avisa de que quien conteste diez preguntas saldría diez veces con su dirección. Olvido pregunta quién va a ver esos teléfonos. Nora solo quiere saber cuántas personas han contestado de verdad. Te encarga la base **EK-2027/BD-v1.0**: se prueba con las encuestas ficticias E01 (martes 1/06) y E02 (miércoles 2/06) y se da por comprobada el viernes 11/06.

❓ ¿Qué tiene que decidir o comprobar Nora: cómo guardar las respuestas, cuántas personas han contestado o las dos cosas?

❓ ¿Qué sabemos de verdad (cuestionarios, fechas, respuestas) y qué suponemos al decir que cada fila es una persona?

❓ ¿Qué prueba con la base permitiría revisar la primera respuesta de Unai antes de cargar los datos?

❓ ¿Qué dato personal hace falta para analizar una encuesta y cuál sobra?

MAPA DE LA UNIDAD

Lo que vamos a ver

11.1 Entidades, claves y relaciones

11.2 Diseño e integridad

11.3 Consultas y búsquedas

11.4 Formularios y validación

11.5 Seguridad, mantenimiento e informes

Tres filas, ¿tres personas?



La consulta devuelve tres entrevistados.

Devuelve tres respuestas de dos personas.





Aclaremos qué representa cada fila.

CLAVE

El JOIN hereda la granularidad de la tabla más detallada: **cada fila es una respuesta**, no una persona. Las personas se cuentan con **COUNT(DISTINCT id_persona)**.

De la hoja de cálculo a la base de datos



Tres tipos de dato

Las **personas** (hogares u oficinas), las **encuestas** que cada una contesta en una fecha y por un canal, y las **respuestas** a cada pregunta. Mezcladas en una hoja, el mismo dato se repite.



Base de datos relacional

Guarda cada dato una sola vez, en tablas que se enlazan por claves. La gestiona un gestor (LibreOffice Base u otro) que entiende **SQL**.



Granularidad

El nivel de detalle de cada registro. Diez filas de Respuestas pueden ser una sola persona: si a «¿qué es una fila?» la respuesta es «depende», la tabla está mal diseñada.



Catálogo de datos

Diccionario EK-2027/DD-v1.0: archivo, tabla, campo, tipo, valores, clave y versión. Base 2027-06-11_estudio_kits_base_v1.0; datos brutos aparte y sin tocar.

CLAVE

Relaciones **uno a varios** (1:N): una persona, varias encuestas; una encuesta, varias respuestas. Las de **varios a varios** (N:M) se resuelven con una tabla intermedia (Pruebas, para personas y kits probados).

Elementos de una base de datos en el estudio de kits

ELEMENTO	QUÉ ES	EN EL ESTUDIO DE KITS
Entidad y tabla	Aquello de lo que se guardan datos	Persona → Personas
Registro y campo	Fila (un caso); columna (un dato)	E01; su fecha
Clave primaria	Identifica cada registro sin repetirse	id_encuesta
Clave foránea	Apunta a la clave primaria de otra tabla	id_persona en Encuestas
Relación	Enlace entre tablas por una clave	Persona 1 → N encuestas
Consulta · formulario · informe	Pregunta · pantalla de entrada · salida	Respuestas por encuesta

Ejemplo: la hoja de la tarjeta de cliente de 2026 tenía una fila por titular; la de la encuesta del reparto, una por cuestionario; la de Unai, una por respuesta. Sumar sus filas como «clientes estudiados» contaría varias veces a la misma persona.

CLAVE

Tabla 2 (catálogo de Encuestas): **id_encuesta**, texto E01, E02..., primaria; **id_persona**, texto P01, P02..., foránea → Personas; **canal**, lista presencial, web, telefono.

11.2 · TABLA 3

La hoja plana propuesta por Unai

Datos ficticios; nombre, dirección y teléfono omitidos.

RESPUESTA	PERSONA, TIPO Y ZONA	ENCUESTA, FECHA Y CANAL	PREGUNTA Y VALOR
R01	P01 · hogar · Norte	E01 · 2027-06-01 · presencial	P1_interes · 4
R02	P01 · hogar · Norte	E01 · 2027-06-01 · presencial	P2_precio · 3
R03	P02 · oficina · Centro	E02 · 2027-06-02 · web	P1_interes · 5

Con solo tres respuestas, P01 y E01 ya salen dos veces. Fallo **de actualización**: si P01 cambia de zona y se olvida una fila, quedan dos zonas. **De inserción**: no se puede dar de alta a quien aún no ha contestado. **De borrado**: al borrar la única respuesta de una encuesta se pierde la encuesta.

CLAVE

Dos personas de prueba y diez preguntas: la hoja plana tendría **20 filas con datos personales**; la base normalizada, 2 filas con ellos (Personas) y 20 sin ellos (Respuestas).

Cada dato, una sola vez

Normalizar es repartir los datos en tablas para guardarlos una vez sin perder ninguna relación. Se comprueba con las tres primeras formas normales:

1

1.ª forma normal

Cada campo guarda un valor y no hay columnas repetidas (P1, P2, P3…).

→ 2

2.ª forma normal

Cada campo depende de toda la clave, no de una parte: con clave id_encuesta + pregunta, el texto dependería solo de pregunta.

→ 3

3.ª forma normal

Ningún campo depende de otro no clave: la zona depende de la persona; el canal, de la encuesta.

CLAVE

La redundancia no solo ocupa: provoca fallos de **actualización, inserción y borrado** y multiplica los datos personales.

11.2 · TABLA 4

Diagrama relacional de EK-2027/BD-v1.0

TABLA (UNA FILA ES…)	CLAVE PRIMARIA	CLAVE FORÁNEA	OTROS CAMPOS
Personas (hogar u oficina)	id_persona	—	tipo, zona
Encuestas (una contestada)	id_encuesta	id_persona	fecha, canal
Respuestas (a una pregunta)	id_respuesta	id_encuesta	pregunta, valor

En SQLite: CREATE TABLE Respuestas (id_respuesta TEXT NOT NULL PRIMARY KEY, id_encuesta TEXT NOT NULL REFERENCES Encuestas (id_encuesta), pregunta TEXT NOT NULL, valor INTEGER NOT NULL CHECK (valor BETWEEN 1 AND 5)); y así Personas (CHECK tipo IN ('hogar', 'oficina')) y Encuestas (CHECK canal IN ('presencial', 'web', 'telefono')).

CLAVE

Personas 1 → N Encuestas → N Respuestas. La clave foránea va siempre en la tabla del lado **«varios»**. Respuestas no lleva id_persona: la persona se obtiene por la encuesta.

🙋 PREGUNTA PARA LA CLASE

¿Dónde se guarda el enlace entre una persona y las encuestas que contesta?

A Una lista de encuestas en un campo de Personas

B id_persona como clave foránea en Encuestas

C id_persona en Respuestas, en cada respuesta

✔ La clave foránea va en la tabla del lado «varios»: id_persona en Encuestas. Una lista en un campo rompe la 1.ª forma normal, y Respuestas obtiene la persona por la encuesta (tabla 4).

Prueba de integridad del martes 8 de junio

Integridad **de entidad**: clave primaria única y no vacía. **Referencial**: toda clave foránea apunta a un registro existente. CHECK y NOT NULL son **reglas de dominio**.

PRUEBA	ENTRADA	ESPERADO	OBTENIDO
Alta válida	R04 → E02, P2_precio, 4	aceptada	aceptada (luego se borra)
Respuesta huérfana	R99 → E09 (no existe)	rechazada	rechazada
Valor fuera de rango	R98 → E01, P1_interes, 7	rechazada	rechazada
Clave repetida	otra R01	rechazada	rechazada
Borrar con dependientes	borrar P01, que tiene E01	rechazada	rechazada

CLAVE

Algunos gestores traen las claves foráneas **desactivadas** (en SQLite, PRAGMA foreign_keys = ON): la base acepta huérfanas sin avisar.

Privacidad desde el diseño



Art. 5.1.c): los datos serán «adecuados, pertinentes y limitados a lo necesario» (minimización). Art. 25.1: medidas apropiadas desde el diseño, «como la seudonimización».

Reglamento (UE) 2016/679 (RGPD), de 27 de abril de 2016 · Consulta la versión vigente

SUJETO OBLIGADO

Quien diseña y trata la base. Para analizar basta con el tipo y la zona: la base no guarda nombre ni dirección. Si hay que volver a llamar, el teléfono va en una tabla aparte de contactos, con acceso solo para quien llama y borrado al terminar el campo.

NO OBLIGA A…

…a guardar el nombre, la dirección y el teléfono en cada fila «por si acaso». El análisis trabaja con P01 y P02, y solo quien tiene los contactos sabe quién es quién.

CLAVE

Un dato **seudonimizado** sigue siendo personal; uno **anonimizado** ya no identifica a nadie.

Una consulta es una pregunta

Selección

Filas que cumplen una condición (WHERE), sin cambiar nada. WHERE canal = 'web' → **E02**.

Agregación

Resume con COUNT, AVG, MIN o MAX y, con GROUP BY, por grupos. AVG(valor) de P1_interes → **4,5** (2 respuestas); COUNT(*) por encuesta → E01: 2 · E02: 1.

Unión (JOIN)

Combina tablas por sus claves: Personas, Encuestas y Respuestas, para saber qué respondió cada persona.

Acción

Cambia datos: INSERT, UPDATE y DELETE. No se deshace con un clic: siempre en una copia.

CLAVE

Búsquedas avanzadas con AND, OR, LIKE 'E0%' o BETWEEN. Se **guardan con nombre** (q_respuestas_por_encuesta) para repetirlas. Antes de ejecutar, escribe qué resultado esperas.

Resultado del JOIN: una fila por respuesta

Personas p JOIN Encuestas e ON e.id_persona = p.id_persona JOIN Respuestas r ON r.id_encuesta = e.id_encuesta

PERSONA	ENCUESTA	RESPUESTA	PREGUNTA Y VALOR
P01	E01	R01	P1_interes · 4
P01	E01	R02	P2_precio · 3
P02	E02	R03	P1_interes · 5

Búsqueda: WHERE canal = 'web' AND fecha BETWEEN '2027-06-01' AND '2027-06-04' → E02. Consulta de control (encuestas sin respuestas): LEFT JOIN Respuestas … WHERE r.id_respuesta IS NULL → ninguna fila. El LEFT JOIN conserva las encuestas sin pareja.

CLAVE

El JOIN hereda la granularidad de la tabla más detallada: tres filas son **tres respuestas**, de dos encuestas y dos personas.

Di siempre qué has contado

COUNT(*) cuenta filas · COUNT(DISTINCT campo), valores distintos

Respuestas	COUNT(*) del JOIN	3
Encuestas	COUNT(DISTINCT e.id_encuesta)	2
Personas	COUNT(DISTINCT p.id_persona)	2
Huérfanas	respuestas sin encuesta	0

Tabla 10 · Recuentos de control de la base (viernes 11/06/2027). La suma por encuesta (2 + 1) coincide con las 3 filas del JOIN.

CLAVE

Tras un JOIN, cuenta personas con **COUNT(DISTINCT id_persona)** y escribe siempre qué has contado.

Siempre en una copia

Prueba del 8 de junio: UPDATE Encuestas SET canal = 'telefono' WHERE id_encuesta = 'E02';

1. **Copia**
   De la base, con fecha.
2. **Selección previa**
   Con la misma condición, para ver qué filas cambian: web antes = 1.
3. **Acción en la copia**
   Se ejecuta el UPDATE en la copia, no en la base de trabajo.
4. **Comprobación**
   Después: web = 0 en la copia. El original sigue con E02 en web.
5. **Anotación**
   En el registro de cambios.

CLAVE

Una acción **no se deshace con un clic**: copia, selección previa, acción en la copia, comprobación y anotación.

Plan de pruebas del formulario (09/06/2027, v1.0)

Selector de encuesta (no se puede escribir E09), lista de preguntas, botones de 1 a 5 («no sabe» y «no procede» se documentan aparte) y campos obligatorios; las encuestas, con un subformulario.

PRUEBA	ENTRADA	ESPERADO	OBTENIDO
Alta válida	E02 · P2_precio · 4	se guarda	se guarda
Dato inválido	E01 · P1_interes · 7	no deja guardar	no deja guardar
Referencia inexistente	encuesta E09	no está en la lista	no está en la lista
Campo vacío	E01 · sin pregunta	mensaje y no guarda	mensaje y no guarda

Un formulario en línea (Google Forms u otro) guarda una hoja plana: se pasa a Encuestas y Respuestas comprobando los recuentos, y la hoja original se guarda sin tocar.

CLAVE

Dos niveles de validación: los datos importados no pasan por el formulario, así que las mismas reglas están en la tabla (CHECK, NOT NULL y clave foránea). Un formulario no está terminado cuando se ve bien, sino cuando rechaza lo que debe rechazar.

Roles, copias y mantenimiento

Principio de **mínimo privilegio**: cada persona tiene solo los permisos que su tarea exige.

ROL	QUIÉN	PUEDE	NO PUEDE
Lectura	Nora, Olvido	Ver informes	Editar nada
Edición	Adrián, Unai	Altas con el formulario	Cambiar tablas, borrar
Administración	Equipo de análisis	Estructura, copias, permisos	—

Con servidor, GRANT y REVOKE (GRANT SELECT ON v_informe TO lectura;); en escritorio, permisos de carpeta e informes de solo lectura, y se prueban. Copia al cierre de cada día de campo y antes de cada cambio de estructura, con fecha y en otro soporte; se conservan las de los últimos cinco días de campo y la de cada versión cerrada. Registro de cambios: v0.9 → v1.0, «se activa la clave foránea de Respuestas».

CLAVE

Una copia **no está comprobada hasta que se restaura** en otra carpeta y da los mismos recuentos (3, 2 y 2). RGPD, art. 32.1.c): «la capacidad de restaurar la disponibilidad y el acceso a los datos personales de forma rápida»; LOPDGDD, art. 5: deber de confidencialidad.

11.5 · TABLA 9

Cada salida, para su destinatario

SALIDA	DESTINATARIO	CAMPOS	FORMATO
Resumen de participación	Nora (dirección)	totales y media de P1 con su base	PDF con versión
Vista de análisis	Equipo de análisis	id_encuesta, tipo, zona, pregunta, valor	CSV UTF-8 con diccionario
Etiquetas de campo	Adrián	id_encuesta, fecha prevista, canal	Hoja de etiquetas

Un informe está vinculado a una consulta guardada: si cambian los datos, se regenera con fecha y versión. Código ICC/ESOMAR (2025): los datos del estudio solo sirven al estudio, los resultados se publican agregados y quien toca la base guarda confidencialidad.

CLAVE

En tu sector. Marketing y publicidad: contar impactos no es contar personas alcanzadas. Gestión de ventas: diez visitas pueden ser un solo cliente.

CASO RESUELTO · PASO A PASO

La base del estudio de kits



El equipo de análisis, para Nora

Viernes 11/06/2027: EK-2027/BD-v1.0, probada con dos encuestas ficticias. Caso ficticio.

1. **Decisión, datos y alcance.** ¿Sustituye la base a las hojas sueltas? P01 (hogar, Norte) y P02 (oficina, Centro); E01 → P01 (01/06, presencial) y E02 → P02 (02/06, web); R01 y R02 → E01, R03 → E02. Alcance: el diseño y las consultas, no el interés del barrio.

2. **Claves y carga.** Tres claves primarias; id_persona en Encuestas e id_encuesta en Respuestas, foráneas, con la integridad referencial activada. Siete registros (2 + 2 + 3).

3. **Consultas y recuentos.** Por encuesta: E01: 2 y E02: 1. JOIN: 3 filas, una por respuesta; COUNT(DISTINCT id_encuesta) = 2; COUNT(DISTINCT id_persona) = 2.

4. **Comprobación.** 2 + 1 = 3 filas del JOIN. Se rechazan R99 → E09, el valor 7, la R01 repetida y el borrado de P01; la copia restaurada da los mismos recuentos. Una media de 4,5 sobre 2 respuestas de prueba no describe a nadie más.

✔ **Dos personas, dos encuestas y tres respuestas.** Nora aprueba EK-2027/BD-v1.0 para el resto del campo, con la tabla de contactos separada y de acceso restringido, y la revisión de recuentos cada viernes.

🙋 PREGUNTA PARA LA CLASE

En la base de Ruedalia (P01, P02; E01 → P01, E02 → P02; R01-R03) Ousmane añade **E03 → P01** con R04 y R05. ¿Qué dan ahora los recuentos?

A 5 respuestas, 3 encuestas y 3 personas

B 5 respuestas, 3 encuestas y 2 personas

C 3 respuestas, 3 encuestas y 2 personas

✔ El JOIN da 5 filas (una por respuesta), 3 encuestas distintas y 2 personas distintas: P01 contesta dos veces. «Ya tenemos tres clientes opinando» es falso (Práctica y autoevaluación 7).

Base relacional comprobada

Con los datos ficticios o autorizados de tu investigación:

1

Diseña

Diagrama relacional con las tablas, qué representa una fila de cada una y sus claves; catálogo con archivo, tabla, campo, tipo, valores, clave y versión.

2

Prueba

Integridad referencial y registro de pruebas (alta válida, huérfana, fuera de rango, borrado con dependientes); consultas de selección, agregación, JOIN con DISTINCT y una acción en copia con el antes y el después.

3

Protege y entrega

Formulario probado; roles, copia restaurada e informe para tu destinatario. Anota la decisión, el responsable, la fecha de revisión y el motivo de cada cambio de versión.

CLAVE

Comprueba tu paso: **cada recuento dice qué cuenta**, y cada prueba tiene su resultado esperado y el obtenido.

APRENDE CON LA IA

La respuesta que recibió Unai

Unai pega en un asistente de IA el resultado del JOIN de la tabla 6 y pregunta cuántas personas han participado. Pulsa cada tarjeta.

«Tres filas tras unir tablas demuestran que hay tres entrevistados»

BULO

Cada fila del JOIN es una respuesta: 3 respuestas de 2 encuestas y 2 personas.

«Para contar personas basta con COUNT(*)»

BULO

COUNT(*) cuenta filas. Las personas se cuentan con COUNT(DISTINCT id_persona) = 2.

«Puedes informar a Nora de que han participado tres personas»

BULO

Han participado 2, y en una prueba de la base: no describe el interés del barrio.

«El asistente puede revisar tu base»

BULO

Solo lee lo que le pegas. No le pegues datos de contacto de personas reales.

REPASO

Ideas clave

- Una base relacional guarda cada dato una vez, en tablas enlazadas por claves.
- Clave primaria: identifica cada registro. Clave foránea: apunta a otra tabla, en el lado «varios».
- Integridad: se rechazan R99 huérfana, el valor 7, otra R01 y el borrado de P01.
- El JOIN da una fila por respuesta: 3 respuestas, 2 encuestas y 2 personas con DISTINCT.
- Formulario y tabla validan a la vez; copia comprobada si se restaura con los mismos recuentos.
- Antes de contar, pregunta qué es una fila: persona, encuesta o respuesta.
- Hoja plana: fallos de actualización, inserción y borrado; 20 filas con datos personales frente a 2.
- Minimización y seudonimización: P01 y P02 en la base; los contactos, aparte y restringidos.
- Consulta de acción: copia, selección previa, acción en la copia, comprobación y anotación.



¡Ahora te toca a ti!

Repasa, practica y comprueba lo que has aprendido.

📝 **Resumen**

✅ **Test**

👆 **Interactivos**

Libro, presentaciones, resúmenes y tests gratis en **josesande.com**