

Asignatura	Datos del alumno	Fecha
Bases de Datos Avanzadas	Apellidos: Condor Saragossin	12/08/2025
	Nombre: Wendy	

Índices en bases de datos

<https://github.com/condorwen27-collab/-ndices-en-bases-de-datos.git>

Tutorial:

https://www.canva.com/design/DAG69CxVHcU/wdH9OLRPEGKHCdUCOO1fg/watch?utm_content=DAG69CxVHcU&utm_campaign=designshare&utm_medium=link2&utm_source=uniquelinks&utllId=h963766de38

Objetivo

Diseñar e implementar un esquema relacional que almacene la información sobre precios de carburantes en España (fuente: Ministerio para la Transición Ecológica y el Reto Demográfico). Realizar la ingesta de los datos desde los ficheros CSV oficiales y validar el correcto funcionamiento mediante consultas.

Herramientas utilizadas

- PostgreSQL 16 (local, instalado en Windows)
- pgAdmin 4 (Query Tool)
- Visual Studio Code (para editar scripts SQL)
- PowerShell (limpieza CSV con Import-Csv / Export-Csv)
- Archivo de datos: preciosEESS_es.csv (descargado del geoportal)

Fases realizadas

Fase 1

Análisis y diseño (E/R):

Se diseñó un modelo E/R con entidades principales:

Provincia, Municipio, Empresa, EstaciónServicio, TipoCarburante, Precio.
 Cardinalidades 1:N según especificación (Provincia→Municipio,
 Municipio→Estación, Empresa→Estación, Estación→Precio, Carburante→Precio).
 Diagrama en draw.io (archivo Modelo ER_Gasolinera.drawio).

Asignatura	Datos del alumno	Fecha
Bases de Datos Avanzadas	Apellidos: Condor Saragossin	12/08/2025
	Nombre: Wendy	

Fase 2

Creación del esquema físico y carga (ETL):

Se creó el DDL (tabla precios_raw + tablas normalizadas).

Se limpió el CSV y se importó a precios_raw.

Se normalizaron datos y se poblaron las tablas finales (provincia, municipio, empresa, tipo_carburante, estacion_servicio, precio).

Se resolvieron errores de datos (coordenadas, textos largos, codificaciones) y se documentó cada corrección.

Esquema relacional (DDL principal)

```
-- provincia
CREATE TABLE provincia (
    id_provincia SERIAL PRIMARY KEY,
    nombre VARCHAR(100) NOT NULL UNIQUE
);

-- municipio
CREATE TABLE municipio (
    id_municipio SERIAL PRIMARY KEY,
    nombre VARCHAR(150) NOT NULL,
    id_provincia INT NOT NULL REFERENCES provincia(id_provincia)
);

-- empresa
CREATE TABLE empresa (
    id_empresa SERIAL PRIMARY KEY,
    nombre TEXT NOT NULL UNIQUE
);

-- tipo_carburante
CREATE TABLE tipo_carburante (
    id_carburante SERIAL PRIMARY KEY,
    nombre VARCHAR(150) NOT NULL UNIQUE
);

-- estacion_servicio
CREATE TABLE estacion_servicio (
    id_estacion SERIAL PRIMARY KEY,
    direccion TEXT,
    codigo_postal TEXT,
```

Asignatura	Datos del alumno	Fecha
Bases de Datos Avanzadas	Apellidos: Condor Saragossin Nombre: Wendy	12/08/2025

```

margen TEXT,
latitud NUMERIC(10,6),
longitud NUMERIC(10,6),
horario TEXT,
tipo_servicio TEXT,
id_empresa INT REFERENCES empresa(id_empresa),
id_municipio INT REFERENCES municipio(id_municipio)
);

-- precio
CREATE TABLE precio (
id_precio SERIAL PRIMARY KEY,
id_estacion INT NOT NULL REFERENCES estacion_servicio(id_estacion),
id_carburante INT NOT NULL REFERENCES tipo_carburante(id_carburante),
precio NUMERIC(10,3),
fecha_actualizacion TIMESTAMP
);

-- precios_raw (tabla temporal/raw)
CREATE TABLE precios_raw (
provincia text,
municipio text,
localidad text,
codigo_postal text,
direccion text,
margen text,
longitud text,
latitud text,
toma_datos text,
precio_gasolina_95_e5 text,
precio_gasolina_95_e10 text,
precio_gasolina_95_e5_premium text,
precio_gasolina_98_e5 text,
precio_gasolina_98_e10 text,
precio_gasoleo_a text,
precio_gasoleo_premium text,
precio_gasoleo_b text,
precio_gasoleo_c text,
precio_bioetanol text,
porc_bioalcohol text,
precio_biodiesel text,
porc_estero_metalico text,
precio_glp text,
precio_gnc text,
precio_gnl text,

```

Asignatura	Datos del alumno	Fecha
Bases de Datos Avanzadas	Apellidos: Condor Saragossin	
	Nombre: Wendy	12/08/2025

```

precio_hidrogeno text,
precio_95_e25 text,
precio_95_e85 text,
precio_adblue text,
precio_diesel_renovable text,
precio_gasolina_renovable text,
precio_metanol text,
precio_amoniacal text,
precio_bgnc text,
precio_bgnl text,
rotulo text,
tipo_venta text,
rem text,
horario text,
tipo_servicio text
);

```

Problemas encontrados y soluciones aplicadas

Comas dentro de campos (dirección): limpieza con Import-Csv | Export-Csv en PowerShell para reescribir correctamente comillas y escapes.

Codificación / caracteres raros (acentos): importar con ENCODING = LATIN1 y/o re-exportar en UTF8.

Coordenadas mal formateadas (ej. 42.178.111): empleé expresiones regulares para validar formato numérico y convertir con REPLACE(...,'.',':')::numeric solo si cumple patrón; valores inválidos devienen NULL.

Campos demasiado largos: cambié columnas problemáticas a tipo TEXT (direccion, horario, tipo_servicio, empresa) para aceptar cualquier longitud.

Asignatura	Datos del alumno	Fecha
Bases de Datos Avanzadas	Apellidos: Condor Saragossin	12/08/2025
	Nombre: Wendy	

JOINS muy estrictos que provocaban 0 resultados: usé TRIM + LOWER para normalizar y matching por dirección + código_postal + municipio (la estrategia B, la más robusta) y, cuando fue necesario, coincidencia flexible con LIKE y unaccent().

Rúbrica

Índices en bases de datos	Descripción	Puntuación máxima (puntos)	Peso %
Criterio 1	Explicación de los índices creados por defecto.	2	20%
Criterio 2	Análisis de rendimiento y creación de índices adicionales para las consultas del laboratorio.	5	50%
Criterio 3	Índices adicionales en las tablas del esquema.	2	20%
Criterio 4	Elaboración de la vídeo memoria obligatoria según las especificaciones, ajustándose a la duración aproximada y cubriendo los aspectos solicitados.	1	10%
		10	100 %