**Caso práctico:** Almacén de datos para el análisis de indicadores de crisis en la eurozona

**Solución PRA1 – Análisis y diseño del *data warehouse***

[**1. Análisis de los requisitos 2**](#_heading=h.m49nibaimlfh)

[**2. Análisis de las fuentes de datos 4**](#_heading=h.dpjyiplb2ma6)

[**Estimación de volumetría 9**](#_heading=h.z03h5ip6oequ)

[**3. Análisis funcional 10**](#_heading=h.bgvdw6oeynf7)

[**4. Diseño del modelo conceptual, lógico  
y físico del almacén de datos 12**](#_heading=h.3zw3ugp6rrt3)

[**Diseño conceptual 12**](#_heading=h.gwzbncfyvqjb)

[**Diseño lógico 14**](#_heading=h.26in1rg)

[**Diseño físico 15**](#_heading=h.lnxbz9)

[**Dimensiones 16**](#_heading=h.m84mk8ogv7fw)

[**Tabla de hechos 18**](#_heading=h.6czh7d1q75pa)

**Nombre estudiante: Lukaz Martin Doehne**

# Análisis de los requisitos

El análisis de los requisitos se basa en identificar las necesidades específicas que tiene una organización particular respecto al análisis de la información. Normalmente, en esta fase, se debe ser previsor y pensar más allá de las necesidades actuales para poder cubrir las futuras.

La necesidad principal de la organización encargada del análisis de indicadores de crisis en la eurozona es disponer de la información integrada para su análisis y su posterior difusión mediante las herramientas de inteligencia de negocio. Estas ayudarán a facilitar la toma de decisiones a todos los usuarios potenciales para garantizar el cumplimiento, entre otros, de los siguientes objetivos:

● Analizar la evolución de la crisis energética de los países de la eurozona.

● Analizar el incremento de precios y su finalidad de consumo.

El diseño de un almacén de datos de un proyecto incluye la creación y la implementación de un modelo dimensional o multidimensional, el diseño y la implementación de procesos ETL y del modelo OLAP y, por último, el diseño de las consultas establecidas en el enunciado.

A continuación, se indica la información necesaria identificada para analizarlo desde diferentes perspectivas:

1. Analizar la evolución de la crisis energética de los países de la eurozona..

● por año

● por año y mes

● por país

● por consumo de energía

● por tipo de energía (gas o electricidad)

1. Analizar el incremento de precios y su finalidad de consumo.

● por año

● por año y mes

● por país

● por finalidad de consumo

Si se tiene en cuenta toda esta información, el sistema podrá responder a múltiples preguntas y, de esta manera, conseguirá cubrir las necesidades de los usuarios potenciales.

A continuación, se indican de manera específica las preguntas que, como mínimo, el sistema debe ser capaz de responder:

● Análisis de la evolución del HICP en la eurozona durante los últimos 5 años.

● Promedio de consumo por país, ordenados por HICP de los últimos 5 años.

● Análisis del top cinco de países con el mayor aumento de HICP del año pasado.

● Análisis del top diez de productos con el mayor incremento de precios durante 2022.

● Análisis del top diez de productos consumidos por país.

● Análisis del top diez de productos consumidos por año.

● Analizar los precios del gas en cada uno de los países de la Unión Europea.

● Analizar los precios de la electricidad en cada uno de los países de la Unión Europea.

# Análisis de las fuentes de datos

En este apartado, se deben revisar las fuentes de datos proporcionadas, qué tipo de información contienen, cuál es su formato y qué datos deben ser cargados. Podéis ver a continuación un análisis detallado para cada tipo de formato.

1) **COICOP.xml**. Relación de todos los códigos y descripción de la finalidad de consumo utilizado por todos los países de la eurozona. Constan de un código alfanumérico y una descripción.

A continuación, se analizan los campos del fichero de texto con etiquetas.

| **Nombre de campo** | **Descripción** | **Tipo** | **Ejemplo** |
| --- | --- | --- | --- |
| Code | Código/Identificador | Texto | CP02122 |
| Description | Descripción del producto | Texto | Wine from other fruits |

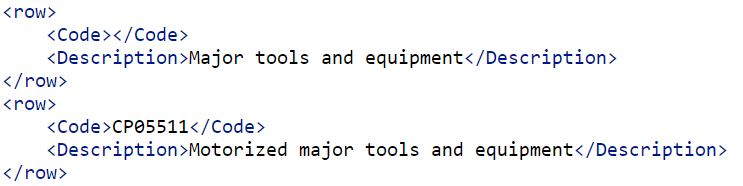
Total de registros: 962.

**Observaciones:**

**●** El campo Code admite valores nulos.

**●** Las descripciones son resúmenes específicos y por lo tanto productos similares no tienen la misma descripción.

A continuación, se muestra una captura del XML para clarificar sus campos:



2) **consumption\_band.csv**. Relación de todos los códigos y la descripción de la franja de consumo de energía. Constan de un código numérico y una descripción de los kWh o GJ consumidos.

A continuación, se analizan los campos del fichero plano separado por comas y con la primera fila con encabezados de cada columna.

| **Nombre de campo** | **Descripción** | **Tipo** | **Ejemplo** |
| --- | --- | --- | --- |
| Code | Código/Identificador | Numérico | 4161901 |
| Description | Franja de consumo de energía | Texto | Band DB: 1000kWh< Consumption  <2500kWh |

Total de registros: 8.

3) **countries.json**. Contiene los nombres de los países en orden alfabético y los elementos de código ISO 3166-1-alpha-2 en formato JSON. La estructura del fichero es la siguiente:

| **Nombre de campo** | **Descripción** | **Tipo** | **Ejemplo** |
| --- | --- | --- | --- |
| name | Nombre de país | Texto | ‘Spain’ |
| code | Código | Texto | ‘ES’ |

Total de registros: 250.

4) **nrg\_pc\_202\_tabular.tsv**. Contiene la evolución del mercado energético europeo de gas, por países. Sus campos incluyen frecuencia, producto, consumo, unidad de medida, impuestos aplicados…

| **Nombre de campo** | **Descripción** | **Tipo** | **Ejemplo** |
| --- | --- | --- | --- |
| freq | Frecuencia con la que se registran los datos (semestral) | Texto | S |
| product | Producto | Numérico | 4100 |
| consom | Consumo de energía (gas) | Numérico | 4141901 |
| unit | Unidad de medida del consumo | Texto | GJ\_GCV |
| tax | Tipo de impuesto aplicado | Texto | I\_TAX |
| currency | Moneda | Texto | EUR |
| geo | Identificador geográfico | Texto | AT |
| TIME\_PERIOD | Período en el que se registraron los datos  (semestral) | Numérico | 20.9300 20.2600  30.4139 |

Total de registros: 1.800.

**Observaciones:**

**●** El campo TIME\_PERIOD viene determinado por semestres desde el primer semestre de 2007 hasta el primer semestre de 2022. Utilizando la convención de año-semestre, como 2007-S1. En total hay 31 campos de TIME\_PERIOD comprendidos entre 2007-S1 hasta 2022-S1. Para simplificar lo muestro cómo una columna, aunque realmente son 31 datos unificados en una columna. Además, el campo incluye carácteres como *e* (*estimated*).

**●** El campo geo viene en formato ISO 3166-1 alpha-2 del país. Es decir, su representación en 2 carácteres.

5) **nrg\_pc\_204\_tabular.tsv**. Contiene la evolución del mercado energético europeo de electricidad, por países. Sus campos incluyen frecuencia, producto, consumo, unidad de medida, impuestos aplicados…

| **Nombre de campo** | **Descripción** | **Tipo** | **Ejemplo** |
| --- | --- | --- | --- |
| freq | Frecuencia con la que se registran los datos (semestral) | Texto | S |
| product | Producto | Numérico | 6000 |
| consom | Consumo de energía (electricidad) | Numérico | 4161901 |
| unit | Unidad de medida del consumo | Texto | KWH |
| tax | Tipo de impuesto aplicado | Texto | I\_TAX |
| currency | Moneda | Texto | EUR |
| geo | Identificador geográfico | Texto | AT |
| TIME\_PERIOD | Período en el que se registraron los datos  (semestral) | Numérico | 20.9300 20.2600  30.4139 |

Total de registros: 1.845.

**Observaciones:**

**●** El campo TIME\_PERIOD viene determinado por semestres desde el primer semestre de 2007 hasta el primer semestre de 2022. Utilizando la convención de año-semestre, como 2007-S1.En total hay 31 campos de TIME\_PERIOD comprendidos entre 2007-S1 hasta 2022-S1. Para simplificar lo muestro cómo una columna, aunque realmente son 31 datos unificados en una columna. Además, el campo incluye carácteres como *e* (*estimated*).

**●** El campo geo viene en formato ISO 3166-1 alpha-2 del país. Es decir, su representación en 2 carácteres.

6) **prc\_hicp\_mv12r\_tabular.tsv**. Contiene el índice de precios armonizados (*Harmonised indices of consumer prices, HICP*). Sus campos incluyen frecuencia, unidad de medida, coicop, geolocalización y período de tiempo.

| **Nombre de campo** | **Descripción** | **Tipo** | **Ejemplo** |
| --- | --- | --- | --- |
| freq | Frecuencia con la que se registran los datos (mensual) | Texto | M |
| unit | Unidad de medida del consumo | Texto | RCH\_MV12MAVR |
| coicop | Código COICOP | Texto | AP |
| geo | Identificador geográfico | Texto | AT |
| TIME\_PERIOD | Período en el que se registraron los datos  (mensual) | Numérico | 20.9300 20.2600  30.4139 |

Total de registros: 14.175.

**Observaciones:**

**●** El campo TIME\_PERIOD viene determinado por meses, a diferencia de los otros dos ficheros .tsv, desde el primer mes de 2020 hasta el primer último mes de 2022. Utilizando la convención de año-mes, como 2010-01.En total hay 36 campos de TIME\_PERIOD comprendidos entre 2020-01 hasta 2022-12. Para simplificar lo muestro cómo una columna, aunque realmente son 36 datos unificados en una columna. Además, el campo incluye carácteres como *d* (*definition differs*).

**●** El campo geo viene en formato ISO 3166-1 alpha-2 del país. Es decir, su representación en 2 carácteres.

# Estimación de volumetría

En los proyectos de diseño de factoría de información corporativa existe una primera fase en la que se realiza una carga inicial y, a posteriori, una segunda fase para realizar las cargas incrementales de los datos nuevos que van llegando.

Una posible estimación del volumen de datos del almacén para la carga inicial de los datos sería la siguiente:

| **Fichero** | **Registros** | **Valores** | **Datos** |
| --- | --- | --- | --- |
| COICOP.xml | 962 | 2 | 1.924 |
| consumption\_band.csv | 8 | 2 | 16 |
| countries.json | 250 | 2 | 500 |
| nrg\_pc\_202\_tabular.tsv | 1.800 | 38 | 68.400 |
| nrg\_pc\_204\_tabular.tsv | 1.845 | 38 | 70.110 |
| prc\_hicp\_mv12r\_tabular.tsv | 14.175 | 40 | 567.000 |
| **Total** | 19.040 | 122 | 707.950 |

# Análisis funcional

A continuación, se propone el tipo de arquitectura para la factoría de información que mejor se adecua al proyecto. Para ello, se consideran los requisitos funcionales y se establece la prioridad entre exigible (E) o deseable (D). En el contexto de esta actividad, los requisitos exigibles son aquellos que se piden en el enunciado, mientras que los deseables son los que complementan la actividad.

Además, en términos de la escala de prioridades, se asigna una prioridad del 1 al 3, en que 1 es completamente prioritario para la actividad y 3 no prioritario.

A continuación, se describen algunos de los requisitos funcionales para el diseño de una factoría de información para la organización, teniendo en cuenta las consideraciones del enunciado:

| **#** | **Requisito** | **Prioridad** | **Exigible/ deseable** |
| --- | --- | --- | --- |
| 1 | Se extraerá de forma adecuada la información de las fuentes de datos. | 1 | E |
| 2 | Se creará un almacén de datos. | 1 | E |
| 3 | Se cargará la información sobre los indicadores de crisis en la eurozona. | 1 | E |
| 4 | Se creará un modelo OLAP para consultas multidimensionales de los usuarios. | 2 | E |
| 5 | Se crearán los informes estáticos solicitados. | 2 | E |
| 6 | Se redactará un manual de carga de datos inicial e incremental. | 3 | D |

En un caso genérico real, se pueden encontrar también otros requerimientos funcionales, como los que se muestran continuación:

● Análisis de viabilidad y análisis de riesgos.

● Creación de procesos de calidad de datos (fuentes de datos).

● Creación de un almacén y un *data mart*.

● Creación de procesos de cargas incrementales.

● Creación de un repositorio de metadatos de gestión del almacén de datos, así como de los procesos ETL, que permita realizar la trazabilidad a lo largo del ciclo de vida de los datos.

● Creación de políticas de seguridad y acceso.

● Implementación de cubos e informes.

● Redacción de manuales de usuario.

Asimismo, dado que estos sistemas frecuentemente forman parte de la implementación de un sistema de inteligencia de negocio, la lista de requerimientos funcionales sería mucho mayor, como puede ser la administración de seguridad en cuanto a datos y a usuarios.

En términos de la arquitectura funcional existen los siguientes elementos:

* Las fuentes de datos de las que se dispone son las siguientes:
  + Un fichero en formato XML.
  + Un fichero en formato CSV, separados por coma (,).
  + Un fichero en formato JSON.
  + 3 ficheros en formato TSV, separados por tabuladores y comas.
* La arquitectura de la factoría de información puede estar formada por varios elementos alojados en la misma máquina:
  + *Staging area* (opcional): cómo tenemos múltiples fuentes de diferente tipo, es conveniente cargarlas y unificarlas en una base de datos intermedia. Un paso que podemos realizar en está etapa es consolidar los datos.
  + *Data Mart* como nos centramos en una única área temática creamos un data mart en lugar de un almacén de datos corporativo.
  + *MOLAP*: a partir de la información de *data mart* crear un cubo multidimensional.

Se podría elegir entre dos diseños para la arquitectura funcional. Para esta actividad propongo la arquitectura funcional que usa un área intermedia (*staging area)*:



Otra arquitectura también podría ser una sin área intermedia (staging area) que identifique las tablas intermedias en el *data mart* con un prefijo en el nombre.



# Diseño del modelo conceptual, lógico y físico del almacén de datos

# Diseño conceptual

Para el correcto desarrollo del DW, es preciso definir los hechos (*facts*), las dimensiones de análisis (*dimensions*), las métricas y los atributos que permitan tener el nivel de granularidad suficiente para la presentación de los objetivos. Estos se han definido en el análisis de los requisitos y de las fuentes de datos.

Del análisis de las fuentes de datos y de los requisitos iniciales, se puede determinar que uno de los hechos que se deben considerar es el siguiente:

● **La evolución de indicadores de la eurozona**. Hace referencia a la información relevante sobre indicadores de la eurozona (HICP).

● **La evolución de la energía**. Hace referencia a la información relevante sobre los precios de la eurozona de gas y/o electricidad.

El **análisis de la evolución de indicadores de la eurozona** determina el diseño de la primera tabla de hechos, como se puede observar a continuación:

| **Tabla de hechos** | **Descripción** |
| --- | --- |
| FACT\_EUROZONE\_INDICATORS | Análisis de los indicadores de la eurozona |

En la siguiente tabla, se indica la métrica de la tabla de hechos FACT\_EUROZONE\_INDICATORS.

| **Métricas** | **Descripción** |
| --- | --- |
| HICP | Harmonised Index of Consumer Prices    Este es un indicador estadístico de precios de consumo de los Estados de la zona del euro, elaborado con los mismos criterios metodológicos en toda la zona, que publica Eurostat y que se utiliza para medir la inflación de los precios de consumo. Está «armonizado» porque todos los países de la Unión Europea siguen la misma metodología. Esto garantiza que los datos de un país puedan compararse con los de otro. |

La métrica de esta tabla de hechos puede ser analizada desde las diferentes perspectivas, a partir de las dimensiones.

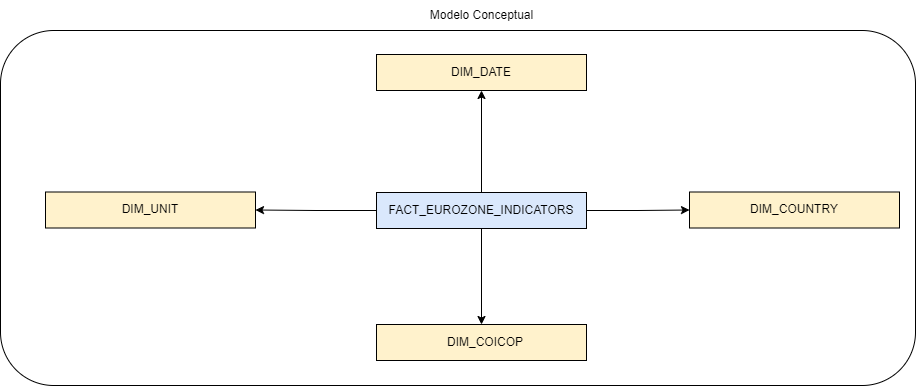
| **Dimensiones** | **Descripción** |
| --- | --- |
| País | País al que hace referencia el precio |
| Fecha | Fecha en la que se registraron los datos |
| Coicop | Finalidad del producto |
| Unidad | Unidad en la que se tomaron los registros |

Podemos optar por 2 tipos de diseño lógico con esta jerarquía:

● El **diseño estrella**. Este modelo parte de una tabla de hechos de la que se desprenden distintas tablas de dimensiones. Las tablas están vinculadas por claves foráneas. Es decir, no se normalizan las bases de datos y se mantienen las jerarquías dentro de una misma tabla.

● El **diseño copo de nieve**. La principal diferencia es que una tabla de dimensiones puede estar interconectada con otra, sin necesidad de recurrir a la tabla de hechos. La creación de nuevas tablas y relaciones, empeora el rendimiento.

El diseño conceptual y sus dimensiones con un **diseño en estrella** es el siguiente:



Este modelo considera las fuentes de datos siguientes:

* Countries.json
* COICOP.xml
* prc\_hicp\_mv12r\_tabular.tsv

Para el análisis de la evolución de la energía se identifica una segunda tabla de hechos; es la siguiente:

| **Tabla de hechos** | **Descripción** |
| --- | --- |
| FACT\_ENERGY | Análisis de los precios de gas/electricidad. |

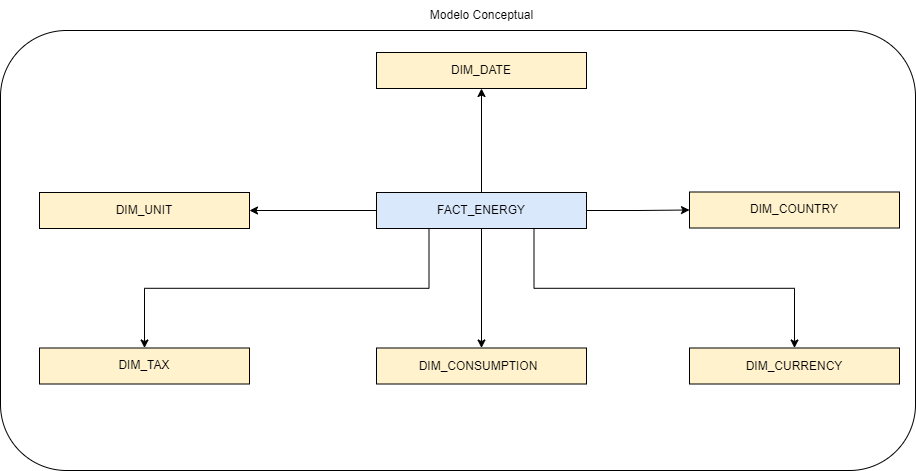
En la siguiente tabla, se indica la métrica de la tabla de hechos FACT\_ENERGY.

| **Métricas** | **Descripción** |
| --- | --- |
| Precio | Precio medio del gas o la electricidad. |
| Producto | Identificador para averiguar si se trata de gas o electricidad. |

Las métricas de esta tabla de hechos puede ser analizada desde las diferentes perspectivas, a partir de las dimensiones.

| **Dimensiones** | **Descripción** |
| --- | --- |
| País | País al que hace referencia el precio de la energía |
| Fecha | Fecha en la que se registraron los datos |
| Unidad | Unidad en la que se tomaron los registros |
| Moneda | Divisa con la que se paga por el servicio |
| Impuestos | Impuestos asociados al consumir la energía |
| Consumo | Franja de consumo media |

El diseño conceptual y sus dimensiones con un **diseño en estrella** es el siguiente:



Este modelo considera las fuentes de datos siguientes:

* Countries.json
* consumption\_band.csv
* nrg\_pc\_202\_tabular.tsv
* nrg\_pc\_204\_tabular.tsv

# Diseño lógico

Una vez obtenido el modelo conceptual del almacén de datos para el análisis de indicadores de crisis en la eurozona, pasamos a realizar su diseño lógico.

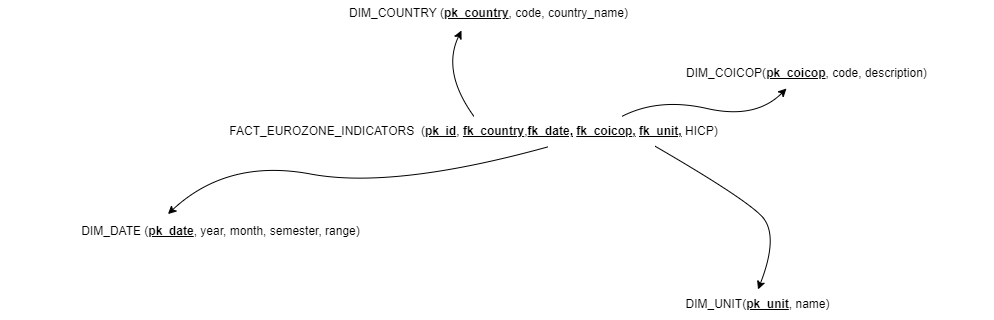
A continuación, se muestra una tabla con las métricas que contiene cada tabla de hechos y que compone el modelo lógico del almacén de datos:

| **Tabla de hechos** | **Métricas** |
| --- | --- |
| FACT\_EUROZONE\_INDICATORS | HICP |
| FACT\_ENERGY | precio |

Después se detallan los atributos de las dimensiones de cada hecho. Específicamente, los atributos de las dimensiones de la tabla de hechos se muestran en la siguiente tabla (FACT\_EUROZONE\_INDICATORS):

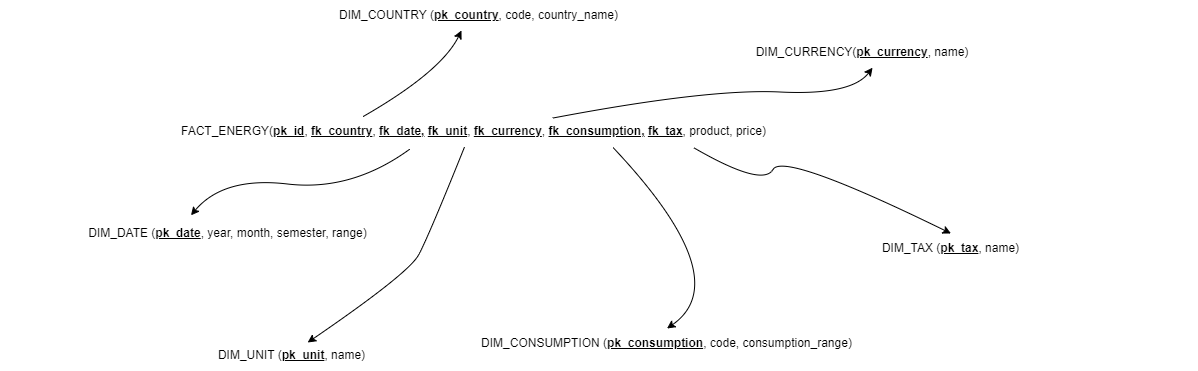
| **Dimensiones** | **Atributos** |
| --- | --- |
| DIM\_COUNTRY | Código, Código del país (2 char.), nombre del país (en inglés) |
| DIM\_DATE | Código, año, mes, semestre, rango (S) |
| DIM\_COICOP | Código, code, descripción |
| DIM\_UNIT | Código, unidad |

La representación visual del **modelo lógico** para el análisis de los indicadores de la eurozona sería la siguiente:



Procedemos igual con el hecho FACT\_ENERGY. Se muestran en la siguiente tabla los atributos de las dimensiones de esta tabla de hechos:

| **Dimensiones** | **Atributos** |
| --- | --- |
| DIM\_COUNTRY | Código, Código del país (2 char.), nombre del país (en inglés) |
| DIM\_DATE | Código, año, mes, semester, rango (M) |
| DIM\_UNIT | Código, unidad |
| DIM\_CURRENCY | Código, divisa |
| DIM\_CONSUMPTION | Código, code, rango (de consumo) |
| DIM\_TAX | Código, impuesto |



# Diseño físico

Para el correcto diseño físico del almacén se deben tener en cuenta los siguientes aspectos:

● El **sistema gestor de bases de datos** con el que vamos a trabajar implementará de una manera concreta los distintos elementos del modelo lógico.

● **El diseño físico**, que debe estar orientado a generar un buen rendimiento en el procesamiento de las consultas.

● La **definición de los procesos de administración** del DW**.**

● La **revisión periódica del diseño físico inicial**, para validar que continúa dando respuesta a las necesidades del cliente.

Una vez que se han determinado las tablas de hechos, las dimensiones, las métricas y los atributos que existen en el modelo, se pueden determinar las claves foráneas que deben definirse en el modelo físico. En este paso, también es necesario tener en cuenta el tamaño adecuado de los atributos (por ejemplo, qué longitud tiene una cadena o si los valores numéricos contienen decimales). También es relevante acordarse de crear correctamente las claves primarias en las dimensiones.

Dado que el modelo de almacén está compuesto por más de una tabla de hechos (*facts*), también se deben revisar las dimensiones que se han definido en el diseño conceptual y en el lógico de cada *fact* y aplicar una visión conjunta del modelo. Esto permitirá definir las dimensiones comunes, como la «fecha» y el «ámbito geográfico», y así simplificar el modelo final y conseguir un rendimiento óptimo en la ejecución de los análisis.

Como es lógico, primero, se crean las tablas de dimensiones y, posteriormente, las tablas de hechos, ya que contienen atributos referenciales a aquellas. De esta forma, se crea cada una de las tablas del almacén de datos.

# Dimensiones

Las dimensiones del modelo podrán estar referenciadas en las tablas de hechos utilizando sus claves primarias, o, en inglés, primary keys (PK). El modelo físico de las dimensiones identificadas es el siguiente:

● **DIM\_Country**: contiene los datos de los países. El primer código es identificativo de la fila, el segundo es un código ISO 3166-1-alpha-2 que representa el país.

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_country (PK)** | Numérico | 8 | 1 |
| code | Texto | 8 | ES |
| country\_name | Texto | 100 | Spain |

● **DIM\_DATE**: contiene la fecha y el rango del registro. Cómo tratamos de diversas representaciones (2022-01, 2022-S1…), guardamos el año seguido del mes/semestre y del rango (M/S).

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_date (PK)** | Numérico | 8 | 1 |
| year | Numérico | 4 | 2022 |
| month | Numérico | 2 | 10 |
| semester | Texto | 2 | S1 |
| range | Texto | 1 | M |

● **DIM\_COICOP**: contiene los datos de la finalidad del producto. El primer código es identificativo de la fila, el segundo es el código COICOP.

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_coicop (PK)** | Numérico | 8 | 1 |
| code | Texto | 8 | AP\_NRG |
| description | Texto | 100 | Wine from other fruits |

● **DIM\_UNIT**: contiene la unidad de energía (gas/electricidad) o del producto.

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_unit (PK)** | Numérico | 8 | 1 |
| name | Texto | 8 | GJ\_GCV |

● **DIM\_CURRENCY**: contiene la divisa del precio.

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_currency(PK)** | Numérico | 8 | 1 |
| name | Texto | 8 | EUR |

● **DIM\_TAX**: contiene el impuesto aplicado al precio.

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_tax(PK)** | Numérico | 8 | 1 |
| name | Texto | 8 | I\_TAX |

● **DIM\_CONSUMPTION**: contiene el rango de energía consumida. El primer código es identificativo de la fila, el segundo es el código del consumo

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_consumption (PK)** | Numérico | 8 | 1 |
| code | Texto | 8 | 4141901 |
| consumption\_range | Texto | 100 | Band DD : 5 000 kWh < Consumption < 15 000 kWh |

# Tabla de hechos

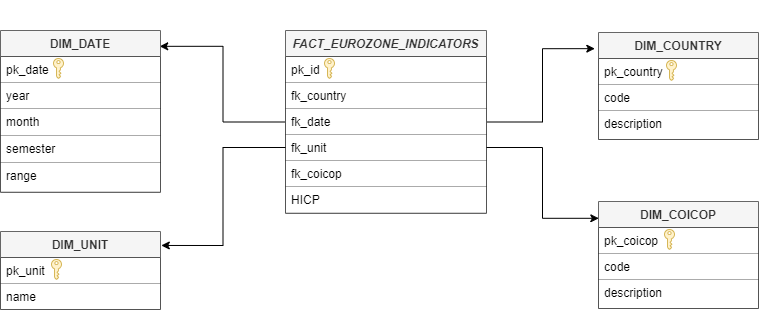
La composición del modelo físico de las tablas de hechos consistirá en la creación de tablas cuyos campos serán las métricas, los atributos y los atributos referenciales definidos en el modelo conceptual y en el modelo lógico. Para crear los atributos referenciales en las tablas de hechos, se definen como claves foráneas las primarias de las dimensiones con las que están relacionadas, siguiendo el diagrama de estrella definido.

El modelo físico de las tablas de hechos del almacén de datos para el análisis de indicadores de crisis en la eurozona está compuesto, entre otras, de las siguientes tablas:

● **FACT\_EUROZONE\_INDICATORS**: es la tabla física que contendrá la información que permitirá realizar el análisis de los indicadores de la eurozona (HICP). Entre otros, tendrá los siguientes campos:

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_id (PK)** | Numérico | 8 | 1 |
| fk\_country (FK) | Numérico | 8 | 231 |
| fk\_date (FK) | Numérico | 8 | 32845 |
| fk\_coicop (FK) | Numérico | 8 | 873 |
| fk\_unit (FK) | Numérico | 8 | 32845 |
| HICP | Numérico | 12,2 | 0.25 |

En la siguiente imagen se muestra el diseño del modelo físico para la tabla de hechos FACT\_EUROZONE\_INDICATORS:



● **FACT\_ENERGY**: es la tabla física que contendrá la información que permitirá realizar el análisis de los precios de energía en la eurozona. Entre otros, tendrá los siguientes campos:

| **Nombre de campo** | **Tipo** | **Tamaño** | **Ejemplo** |
| --- | --- | --- | --- |
| **pk\_id (PK)** | Numérico | 8 | 1 |
| fk\_country (FK) | Numérico | 8 | 231 |
| fk\_date (FK) | Numérico | 8 | 32845 |
| fk\_currency (FK) | Numérico | 8 | 873 |
| fk\_unit (FK) | Numérico | 8 | 32845 |
| fk\_consumption (FK) | Numérico | 8 | 4323 |
| fk\_tax (FK) | Numérico | 8 | 982 |
| product | Numérico | 8 | 4100 |
| price | Numérico | 12,4 | 0.0925 |

En la siguiente imagen se muestra el diseño del modelo físico para la tabla de hechos FACT\_ENERGY:

