

TECNOLOGÍAS DE INTELIGENCIA DE NEGOCIO



UCA

Universidad
de Cádiz

MEMORIA DE PROYECTO

Autores:

Francisco Javier Molina Rojas
Francisco Mercado Espinosa
Jesús Gabriel Gómez Tocino

FEBRERO 2024

Índice

	Página
I Hito 1: Conceptualización del proyecto	2
1. Motivación	2
2. Objetivos	2
3. Preguntas	2
4. Métricas a Analizar	2
5. Características de los Sistemas de Origen	3
6. Características del Modelo de Datos	3
II Hito 2: Arquitectura del DW/DM	4
7. Modelo de datos diseñado para el DW y/o los DMs	4
7.1. Esquema estrella	4
7.2. Tabla Educacion_Espana_Hechos	4
7.3. Tabla D_CCAA	4
7.4. Tabla D_Educacion	4
7.5. Tabla D_Estudiante	4
8. Descripción del esquema OLAP	5
8.1. Dimensión Comunidad Autónoma (C.C.A.A.)	5
8.2. Dimensión Educación	5
8.3. Dimensión Estudiante	5

Parte I

Hito 1: Conceptualización del proyecto

1. Motivación

La educación es un pilar fundamental de cualquier sociedad, y su análisis detallado es crucial para comprender y mejorar el sistema educativo. Con la creciente cantidad de datos disponibles, es esencial emplear soluciones de Inteligencia de Negocio (BI) para desentrañar patrones y tendencias ocultas. Este proyecto busca aprovechar los vastos datos educativos de España para obtener *insights* significativos que contribuyan a una educación más efectiva y equitativa, adaptándose a las necesidades cambiantes de los estudiantes y la sociedad.

2. Objetivos

Los principales objetivos de nuestra solución BI son:

- Analizar la distribución de recursos y su impacto en la calidad educativa.
- Comprender las tendencias demográficas y su influencia en la educación.
- Identificar áreas de mejora en la infraestructura educativa y la asignación de recursos.
- Evaluar el rendimiento académico en diferentes niveles y modalidades educativas.
- Promover una educación equitativa y accesible en todas las comunidades autónomas.

3. Preguntas

La solución BI busca responder preguntas como:

1. ¿Cómo se distribuye el gasto público en educación entre las diferentes comunidades autónomas y niveles educativos?
2. ¿Cuál es la correlación entre la inversión en educación y los resultados académicos?
3. ¿Cómo afectan factores demográficos como el sexo, la edad y la nacionalidad a la educación?
4. ¿Cuáles son las tendencias en cuanto a abandono escolar y éxito académico?
5. ¿Cómo influye la infraestructura tecnológica en los resultados educativos?

4. Métricas a Analizar

Las principales métricas a analizar incluyen:

- Gasto público y privado en educación.
- Ratio de profesores-alumnos en diferentes niveles educativos.
- Tasa de rendimiento, éxito y abandono escolar.
- Distribución demográfica de estudiantes y profesores.
- Uso de tecnologías y recursos en la educación.

5. Características de los Sistemas de Origen

Los sistemas de origen de los datos son archivos XLSX extraídos del portal de datos “EducaBase” del Ministerio de Educación español. Estos archivos contienen datos detallados sobre diversos aspectos del sistema educativo español, desde gastos y recursos hasta rendimiento académico y demografía estudiantil y docente.

6. Características del Modelo de Datos

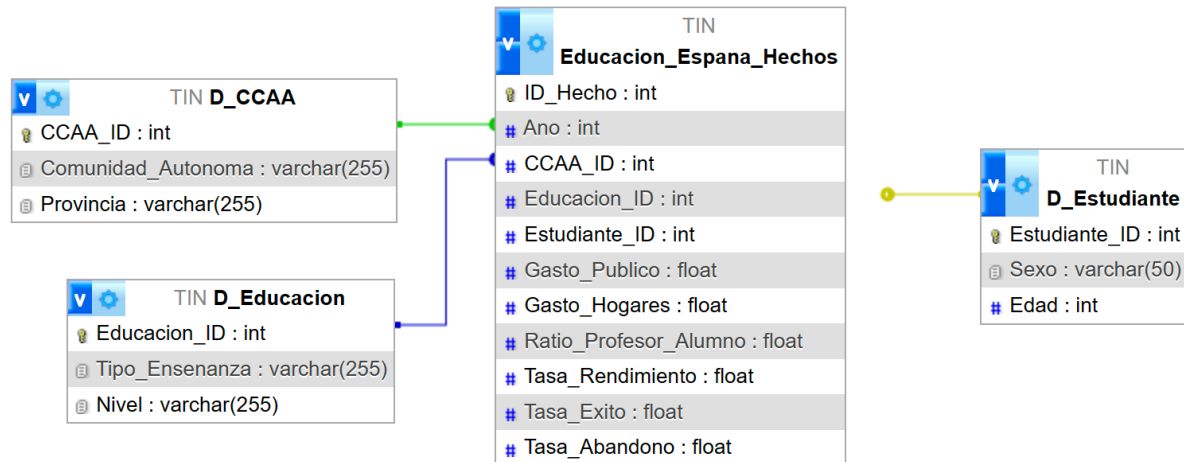
El modelo de datos se basa en un esquema estrella centrado en una tabla de hechos, que almacena métricas clave, y varias tablas de dimensiones que proporcionan contextos detallados (temporal, geográfico, educativo, y demográfico). Este enfoque facilita la realización de análisis multidimensionales y ayuda a descubrir *insights* relevantes sobre el sistema educativo español.

Parte II

Hito 2: Arquitectura del DW/DM

7. Modelo de datos diseñado para el DW y/o los DMs

7.1. Esquema estrella



7.2. Tabla Educacion_Espana_Hechos

Tabla de hechos. Pieza central del esquema, almacena información relativa todos los esquemas.

7.3. Tabla D_CCAA

Tabla de dimensión de comunidades autónomas, almacena información relativa a las comunidades autónomas. Podrá ser usada como dimensión al consultar los hechos.

7.4. Tabla D_Educacion

Tabla de dimensión de educación, almacena información relativa a los distintos niveles educativos. Podrá ser usada como dimensión al consultar los hechos.

7.5. Tabla D_Estudiante

Tabla de dimensión de estudiante, almacena información relativa a las personas que estudian. Podrá ser usada como dimensión al consultar los hechos.

8. Descripción del esquema OLAP

El esquema OLAP está directamente relacionado con las dimensiones escogidas para el proyecto.

8.1. Dimensión Comunidad Autónoma (C.C.A.A.)

- *CCAA_ID*: Clave primaria que identifica a cada registro de la dimensión.
- *Comunidad_Autonomas*
- *Provincia*: Provincia de una determinada C.C.A.A. (si procede).

8.2. Dimensión Educación

- *Educacion_ID*: Clave primaria que identifica a cada registro de la dimensión.
- *Tipo_Ensenanza*: Determina si la enseñanza es de tipo pública, privada o concertada.
- *Nivel*: Primaria, Secundaria, Grado Medio, Bachillerato, Enseñanza Superior (Grado Superior, Grado Universitario).

8.3. Dimensión Estudiante

- *Estudiante_ID*: Clave primaria que identifica a cada registro de la dimensión.
- *Sexo*: Mujer, Hombre, No especificado
- *Edad*