TECHNOLOGY RAIN JOURNAL

ISSN 2953-464X | Vol. 5, Núm. 1 (2026)

Artículo de Investigación Original

Implementación de un Data Warehouse para la toma de decisiones en el proyecto de investigación de la evaluación de campos electromagnéticos de radio, televisión y celulares en la ESPOCH

Implementation of a Data Warehouse for Decision-Making in the Research Project on the Evaluation of Electromagnetic Fields from Radio, Television, and Mobile Phones at ESPOCH

Danilo Llangarí Silva[1] **Pedro Severo Infante Moreira**[2]

[1]ESPOCH, Facultad de Informática y Electrónica, Escuela de Posgrado, Ecuador;
danilo.llangari@espoch.edu.ec; ORCID: 0009-0008-6960-0115

[2]ESPOCH, Facultad de Informática y Electrónica, Telecomunicaciones, Ecuador; pinfante@espoch.edu.ec;
ORCID: 0000-0003-0160-5880

Recibido: 08/10/2025 **Aceptado:** 04/12/2025 **Publicado:** 20/04/2026

DOI: https://doi.org/10.55204/trj.v5i1.e164

Resumen. El incremento en el uso de tecnologías de radiodifusión y telecomunicaciones ha incrementado la exposición de la población a campos electromagnéticos, generando grandes volúmenes de datos en el marco del proyecto de evaluación desarrollado en la ESPOCH. Las campañas de medición produjeron alrededor de 13.000 archivos CSV con más de 1.5 millones de registros geoespaciales y espectrales, lo cual dificulta su integración, depuración y análisis mediante herramientas tradicionales. Esta investigación aborda la necesidad de integrar y analizar dicha información mediante la implementación de un Data Warehouse orientado a la toma de decisiones, fundamentado en la metodología Kimball, con un modelo dimensional en estrella y un proceso ETL programático en Python, comparado con una herramienta gráfica. El caso de estudio demuestra que el enfoque programático ofrece mejor rendimiento, además de facilitar la flexibilidad y la integración de datos espaciales y temporales. Se concluye que el DW constituye una solución robusta y escalable para apoyar decisiones basadas en evidencia en contextos académicos y técnicos.

Palabras clave: Data Warehouse; DW; Procesos ETL; Python; Kimball; Campos electromagnéticos.

Abstract. The increasing use of broadcasting and telecommunications technologies has heightened population exposure to electromagnetic fields, generating large volumes of data within the evaluation project conducted at ESPOCH. The measurement campaigns produced approximately 13,000 CSV files containing more than 1.5 million geospatial and spectral records, which poses significant challenges for integration, cleansing, and analysis using traditional tools. This study addresses this need through the implementation of a Data Warehouse designed to support decision-making, built following the Kimball methodology, using a star-schema dimensional model and a programmatic ETL process in Python, benchmarked against a graphical tool. The case study shows that the programmatic approach delivers better performance while also enabling greater flexibility and more effective integration of spatial and temporal data. The results indicate that the Data Warehouse represents a robust and scalable solution for supporting evidence-based decision-making in academic and technical contexts.

Keywords: Data Warehouse; DW; ETL Processes; Python; Kimball; Electromagnetic Fields.
Licencia: Creative Commons Atribución 4.0 Internacional (CC BY 4.0).

INTRODUCCIÓN

Actualmente, el creciente uso de tecnologías que permiten mayor interconexión entre individuos tales como radio, televisión y telecomunicaciones celulares, ha incrementado significativamente la exposición de la población a campos electromagnéticos (CEM), fenómeno que ha sido objeto de estudio por parte de la academia respecto a los posibles efectos de tal exposición sobre la salud humana, generando una cantidad significativa de datos provenientes de múltiples fuentes que puede conllevar a información duplicada o que presente inconsistencias.

En el marco del proyecto de Investigación de evaluación de campos electromagnéticos en la ESPOCH, las campañas de medición realizadas con el equipo Narda SRM-3006 generaron un volumen considerable de información, compuesto por aproximadamente 13.000 archivos CSV, que en conjunto superan los 1.5 millones de registros individuales de mediciones geoespaciales y espectrales. Este volumen masivo de datos, excede la capacidad operativa de los flujos tradicionales de gestión basados únicamente en herramientas de escritorio y análisis manual. La manipulación eficiente de esta cantidad de información resulta crítica para garantizar la calidad, consistencia y trazabilidad de los datos utilizados en los estudios de exposición poblacional, y demanda soluciones que permitan integrar y transformar grandes conjuntos de datos de manera sistemática.

En consecuencia, es un desafío fundamental la gestión eficiente de grandes volúmenes de datos para optimizar procesos. La correcta gestión y análisis de estos datos es crucial para garantizar la seguridad, el cumplimiento normativo y la toma de decisiones informadas en el proyecto de investigación. En este contexto los Data Warehouse se han convertido en una herramienta esencial en proyectos donde se requiere correlacionar información espacial, temporal y espectral, como ocurre en la evaluación de CEM.

Un Data Warehouse (DW) se define como un repositorio centralizado diseñado para almacenar datos estructurados de manera optimizada para análisis y consultas.

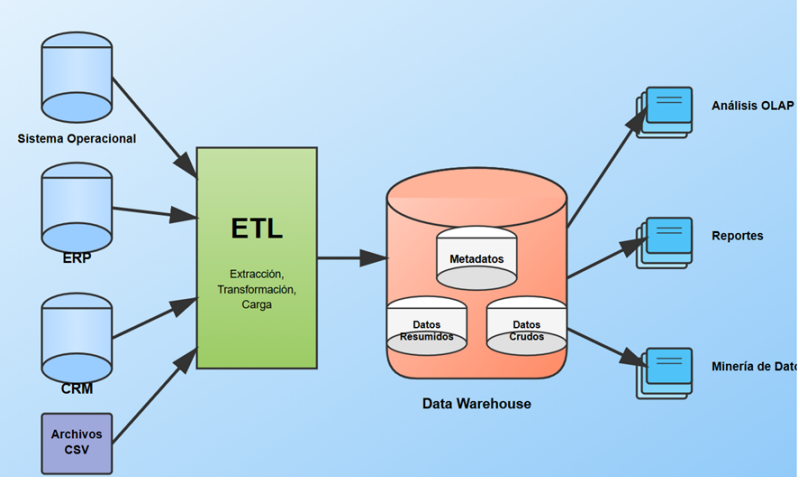
Esta concepción es coherente con la definición de Inmon (2005), quien lo caracteriza como “una colección de datos orientada a un tema, integrada, no volátil y variable en el tiempo, destinada a apoyar la toma de decisiones gerenciales”. Bill Inmon fue uno de los primeros autores que escribieron sobre el tema de los DW.

De igual forma se ha establecido que el Data Warehousing es el proceso mediante el cual los datos son transformados y agrupados dentro de un almacén de datos, permitiendo una fusión de los mismos los cuales provienen de diversas fuentes hacia un solo contexto, lo cual permite un acceso rápido a la información y por consiguiente a la toma de decisiones (Silva et al., 2021). Como se observa en la Figura 1.

Bajo este enfoque, y debido a la escalabilidad proyectada en la investigación y los datos de las campañas de medición proyectadas a futuro tanto temporales y espaciales, generan miles de archivos lo que nos lleva a considerar que el DW puede evolucionar hacia Big Data si el volumen crece.

Y tal como afirman Santoso y Yulia: Los datos contenidos en el DW se pueden utilizar como

Figura 1: Arquitectura de un Data Warehouse (Dragomir, 2016).



entrada para el sistema de aplicación, por ejemplo, para un dashboard... Con la llegada del big data, los almacenes de datos tradicionales no pueden gestionar grandes cantidades de datos... La enorme carga de trabajo, los usuarios simultáneos y los volúmenes de datos requieren la optimización del diseño lógico y físico. Por lo tanto, el procesamiento de datos debe realizarse en paralelo.

Dado que la gestión de volúmenes crecientes de datos exige estructuras flexibles y orientadas al análisis, es necesario apoyarse en metodologías que faciliten la construcción modular y escalable del DW, los autores indican que "La metodología de Kimball, aplicada ampliamente en organizaciones, se presenta como una alternativa sencilla y rápida, para la construcción de un almacén de datos de manera incremental" (Ñauta & Silva, 2025)

No obstante, un modelo dimensional bien estructurado resulta insuficiente si los datos que lo alimentan no mantienen estándares adecuados de calidad; por ello, las etapas ETL constituyen un componente determinante en la confiabilidad de los resultados. En este contexto, el análisis de los autores sugiere que la calidad de los datos transformados durante las etapas ETL impacta directamente en la precisión de los resultados analíticos, ya que indica que la precisión y relevancia de Business Intelligence & Analytics (BI&A) dependen de la capacidad de llevar datos de alta calidad al almacén de datos desde fuentes internas y externas utilizando el proceso ETL. (Souibgui, Atigui, Zammali, Cherfi & Ben Yahia. 2019). Un aspecto que es especialmente relevante para estudios basados en mediciones geoespaciales y temporales.

Tomando en cuenta los desafíos identificados en cuanto a volumen, calidad y procesamiento eficiente de la información, la presente investigación tiene como objetivo la Implementación de un Data Warehouse para la toma de decisiones en el proyecto de investigación de la evaluación de campos electromagnéticos de radio, televisión y celulares en la ESPOCH, considerando su aplicabilidad en un modelo de datos en estrella previamente definido y la ejecución de un proceso ETL utilizando Python comparando su desempeño con soluciones gráficas, con el fin de evidenciar el valor agregado que aporta esta elección tecnológica, por lo que contribuye también un modelo replicable para proyectos en contextos similares.

METODOLOGÍA

El diseño de un DW no solo implica almacenar datos, sino también estructurarlos bajo metodologías que aseguren consistencia y escalabilidad.

La presente investigación se desarrolló bajo un enfoque cuantitativo, orientado a medir el rendimiento de diferentes procesos ETL y el comportamiento del Data Warehouse implementado. El estudio corresponde a una investigación aplicada y experimental, dado que se diseña, implementa y evalúa una solución tecnológica en un entorno real utilizando los datos completos generados por el proyecto institucional de mediciones de campos electromagnéticos.

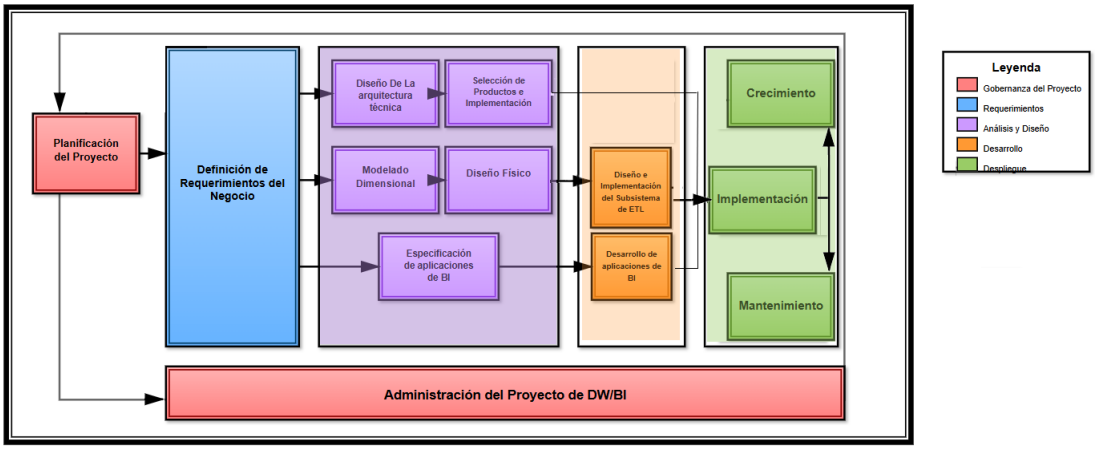
El proyecto completo generó 12.992 archivos CSV, obtenidos de las campañas de medición con el equipo Narda SRM-3006, que en conjunto contienen más de 1.500.000 registros individuales de mediciones electromagnéticas. Estos archivos constituyen la población total de datos del estudio y fueron utilizados para el diseño del modelo dimensional, la construcción del Data Warehouse y la validación de procesos integrales de carga y transformación.

2.1 Kimball

La metodología de Kimball, llamada Modelo Dimensional (Dimensional Modeling), se basa en lo que se denomina Ciclo de Vida Dimensional del Negocio (Business Dimensional Lifecycle). Esta metodología es considerada una de las técnicas favoritas a la hora de construir un Data Warehouse. (Silva et al., 2021)

La metodología de Ralph Kimball, propone una serie de fases para el diseño, desarrollo e implementación de un Data Warehouse. Estas etapas incluyen planificación del proyecto, definición de requerimientos del negocio, diseño dimensional, construcción de procesos ETL, e implementación y mantenimiento de la solución de BI (Kimball et al., 2008), como se muestra en la Figura 2.

Figura 2: Fases de metodología de Ralph Kimball (Kimball et al., 2008).



El enfoque de la investigación se fundamenta en un diseño bajo la metodología Kimball que garantizan la integración y explotación de información como apoyo a las organizaciones para tomar mejores decisiones en cuanto a la inteligencia de negocios ajustándose a las necesidades empleando normativas y planificaciones que facilitan el sistema de almacenamiento de datos, además autores recomiendan la metodología Kimball como la más adecuada para proyectos pequeños, para realizar

un mejor uso de la información y permitir la integración de BI en las actividades estratégicas, tácticas y operacionales (Guerrero & Rodas, 2022).

En el Modelo Dimensional se constituyen modelos de tablas y relaciones con el propósito de optimizar la toma de decisiones, a partir de consultas hechas en una base de datos relacional que están ligadas con la medición o un conjunto de mediciones de los resultados de los procesos de negocio. (Silva et al., 2021)

Asimismo, la adopción de un enfoque metodológico Kimball permite estructurar fases claras del proyecto requerimientos, modelado dimensional, esquema estrella, subsistema ETL, aplicaciones de BI.

El estudio de Ñauta y Silva lo sintetiza así: "La Metodología de Kimball debe cumplir ciertas fases antes de obtener los resultados esperados entre las más relevantes están requerimientos iniciales, modelado dimensional, proceso ETL, despliegue y generación de informes."

De forma complementaria, experiencias recientes muestran el uso de Pentaho Data Integration (PDI) y el diseño Star Schema (hechos y dimensiones) para optimizar el análisis.

2.2 Esquema estrella

El esquema en estrella consta básicamente de tablas de hechos y tablas de dimensiones. Cada tabla de hechos se compone además de claves foráneas de diversas dimensiones y medidas, así como de dimensiones degeneradas, si las hubiera. (Kirmani, 2017)

Los resultados del proceso de gestión de datos y la implementación de un DW con esquema en estrella demuestran que puede ayudar a depurar e integrar los datos en un mejor formato. El DW mejora la estructura y la consistencia de los datos, lo que facilita su análisis y gestión. El esquema en estrella también ayuda a analizar los datos para resolver problemas. (Wijaya, Wiratama & Wijaya, 2024)

2.3 ETL

Un componente crítico en la implementación del DW es el proceso ETL (Extract, Transform, Load), y la implementación constituye un componente esencial en la construcción de DW, ya que permite integrar información heterogénea y garantizar su calidad para análisis posteriores. Por lo que la elección de herramientas para desarrollar dichos procesos impacta directamente en la productividad, flexibilidad y rendimiento del sistema.

El lenguaje Python se ha consolidado como una alternativa robusta frente a soluciones tradicionales basadas en entornos gráficos. Tal como se afirma: Python está predominantemente orientado a objetos, pero hasta cierto punto también admite la programación funcional, por ejemplo, al permitir el uso de funciones o expresiones lambda como argumentos. (Jensen et al., 2021).

Esta característica, sumada a su ecosistema de librerías, facilita la creación de flujos ETL programáticos que aprovechan la expresividad del código y la integración con otras tecnologías. En contraste, herramientas como Pentaho Data Integration (PDI) presentan limitaciones en términos de rapidez y simplicidad en el desarrollo. Al respecto los autores presentan gráficas de tasas de carga y utilización de recursos en comparación con pygrametl, framework que promueve una aproximación programática en Python, permitiendo a los desarrolladores construir y testear flujos ETL mediante

código, lo cual puede ofrecer mayor flexibilidad y productividad (Jensen et al., 2021)

En la selección de la herramienta óptima para esta investigación se realizó una comparativa con una herramienta de interfaz gráfica Pentaho, valorando los criterios más comunes de análisis de acuerdo a como se muestra en la tabla 1.

Tabla 1. Criterios para valoración.

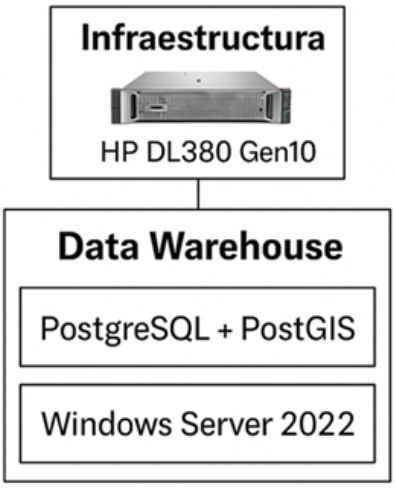
Criterio	Descripción	Métrica / Unidad	Herramienta
Tiempo de ejecución	Duración total del proceso ETL	Segundos (s)	Time o perf_counter en Python; logs en Pentaho
Uso de CPU	Porcentaje promedio de CPU utilizado durante la ejecución	%	Psutil en Python; monitor del sistema para Pentaho
Uso de memoria	Memoria RAM consumida por el proceso ETL	MB	Psutil en Python; monitor del sistema para Pentaho
Escalabilidad	Comportamiento ante incremento de volumen de datos (1 día, 7 días, 28 días, 24 archivos dia)	Escala cualitativa (Alta/Media/Baja)	Esfuerzo
Flexibilidad	Capacidad para integrar cálculos personalizados y librerías externas	Escala cualitativa (Alta/Media/Baja)	Manejo nulos, formatos

En la línea de ETL programático con Python, se destaca el papel de Pandas para tareas de lectura/escritura y preparación de datos (nulos, indexación, agrupaciones, fusiones y codificación), tal como se señala, la analítica de datos en Python inició un crecimiento vertiginoso a partir del desarrollo de la librería Pandas, que otorga funcionalidades de simple implementación y gran utilidad para el procesamiento de datos crudos. La librería Pandas ha impulsado el crecimiento de la analítica de datos en Python al ofrecer funcionalidades de implementación sencilla y gran utilidad para el procesamiento de datos crudos, lo cual constituye una de las tareas más laboriosas del proceso ETL. (Kamlofsky, Manzano & Lopez Yse, 2024).

Si bien muchas arquitecturas de Data Warehouse usan motores especializados (como sistemas MPP o columnar), también es factible usar PostgreSQL como motor de base de datos para un DW, especialmente en plataformas de costo bajo o medianas. El estudio de viabilidad de una plataforma de bajo costo para Data Warehouse demuestra que PostgreSQL puede utilizarse como base para cargas analíticas, aunque con consideraciones de rendimiento específicas para entornos OLAP y grandes volúmenes: la ejecución de consultas complejas con múltiples agregaciones y joins es parte de la carga de trabajo del DW (Almeida, 2011), como se aprecia a la en la figura 3.

RESULTADOS Y DISCUSIÓN

Figura 3: Infraestructura del Data Warehouse para el análisis de campos electromagnéticos.



3.1 Resultados de la Implementación del Data Warehouse

La arquitectura propuesta se implementó exitosamente en PostgreSQL con extensión PostGIS, siguiendo un modelo dimensional en estrella, como se muestra en la figura 4. Este modelo incluye tablas de hechos para mediciones temporales y espaciales, factores de extrapolación y valores RMS por nodo, así como dimensiones para tiempo, nodo, servicio/frecuencia, equipo, estándar, campaña, modelo virtual y tipo de campo.

Figura 4: Modelo Estrella.



El diseño permite integrar datos heterogéneos provenientes de los archivos CSV del equipo Narda SRM-3006 y soporta consultas analíticas rápidas, además de integración con ArcGIS para la generación de mapas electromagnéticos.

3.2 Procesos ETL en Python

Se desarrolló un proceso ETL en Python utilizando principalmente la librería Pandas, que permitió la extracción, transformación y carga de datos de forma flexible y eficiente de acuerdo a las particularidades de los archivos. Las transformaciones incluyen la conversión de coordenadas DMS a decimales, normalización de datos, y cálculo de métricas adicionales como factores de extrapolación, Media

Cuadrática, Promedios diarios, semanales y mensuales entre otros procesos propios del proyecto de investigación, como se puede observar en la Figura 5.

Esta perspectiva programática demostró ventajas significativas frente a herramientas tradicionales, como Pentaho, en términos de personalización y control del flujo de datos, tal como lo señalan (Kamlofsky et al., 2024), quienes destacan que el uso de Pandas simplifica tareas ETL y reduce la complejidad del preprocesamiento.

Figura 5: Fragmento Script Python.

```
    dec = deg + minutes/60.0 + seconds/3600.0
    if hemi in ['S','W']: dec = -dec
    return dec

def parse_narda_csv(file_path, campaign_type='Temporal', node_code='NODO_001',
    node_name='Nodo ESPOCH 001', modelo_virtual=None):
    with open(file_path, 'r', encoding='utf-8', errors='ignore') as f:
        lines = [ln.strip() for ln in f.readlines()]

    # Separar cabecera vs tabla de servicios
    header, table_lines, in_table = [], [], False
    for ln in lines:
        if ln.startswith(';Fmin'):
            in_table = True
            table_lines.append(ln)
            continue
        if not in_table: header.append(ln)
        else:
            if ln: table_lines.append(ln)

    # Cabecera como pares clave;valor
    meta = {}
    for ln in header:
        if ';' in ln:
            k, v = ln.split(';', 1)
            meta[k.strip()] = v.strip()

    # Metadatos
    date_str, time_str = meta.get('Date'), meta.get('Time')
    dt = None
    if date_str and time_str:
        try:
            dt = datetime.strptime(f"{date_str} {time_str}", "%d.%m.%Y %H:%M:%S")
        except Exception:
            dt = None
    minuto_del_dia = dt.hour*60 + dt.minute if dt else None

```

3.3 Comparación Python vs Pentaho

Para examinar el rendimiento, se ejecutaron pruebas con un subconjunto de 240 archivos CSV de los obtenidos en las mediciones de la campaña Temporal, en el proceso que se observa en la figura 6. Esta muestra operativa se utilizó exclusivamente para el benchmarking de tiempos, uso de CPU y memoria.

Sin embargo, el proceso ETL final en Python fue validado y ejecutado sobre la totalidad de los 12.992 archivos, que en conjunto contienen más de 1.5 millones de registros, garantizando así la escalabilidad y validez del Data Warehouse implementado.

Los hallazgos mostraron que el proceso ETL en Python presentó un menor tiempo de ejecución y menor consumo de recursos frente a Pentaho, del mismo modo Python nos permite integrar los procesos que en el proyecto de investigación se realizan con un software de cálculo numérico.

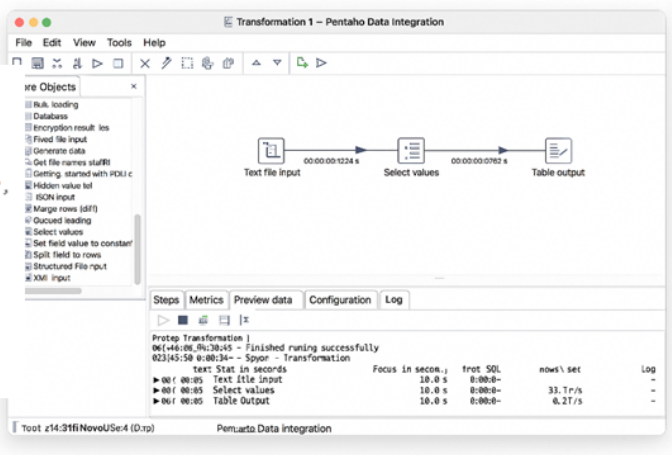
Figura 6: ETL Python vs ETL Pentaho.

```
import time, psutil
start_time = time.perf_counter()
process = psutil.Process()

# Ejecutar ETL
run_batch(input_folder="csv_temporales", campaign_type="Temporal",
node_code="NODO_ESPOCH_001", node_name="Nodo ESPOCH 001")

end_time = time.perf_counter()
elapsed = end_time - start_time
cpu_usage = process.cpu_percent(interval=None)
mem_usage = process.memory_info().rss / (1024 * 1024)

print(f"Tiempo total: {elapsed:.2f} s")
```



Estas observaciones coinciden con estudios previos que resaltan la eficiencia de soluciones programáticas frente a herramientas gráficas (Kamlofsky et al., 2024). Los mismos que se pueden apreciar en la tabla 2.

Tabla 2. Resultados de comparación.

Métrica	Python ETL	Pentaho ETL
Tiempo (s)	120	180
CPU (%)	35	50
Memoria (MB)	2.5	4
Escalabilidad	Alta	Media
Flexibilidad	Alta	Baja

Calidad de Datos

Además del rendimiento, se consideró la calidad de los datos como un factor crítico para garantizar la confiabilidad del análisis. En este sentido, Barreno & Silva (2025) enfatizan la importancia de aplicar estándares como ISO/IEC 25012 y metodologías como TDQM para evaluar dimensiones clave como unicidad, completitud, conformidad y consistencia. Si bien el presente estudio no implementó un sistema formal de medición de calidad, se incorporaron validaciones básicas en el ETL para asegurar la integridad de los datos, tales como control de duplicados y normalización de formatos. Este aspecto se plantea como una línea futura de mejora, integrando herramientas como Great Expectations para automatizar pruebas de calidad.

3.5 Discusión Crítica

Los resultados confirman que la implementación del DW y el uso de Python para ETL aportan beneficios significativos en términos de rendimiento, flexibilidad y capacidad de integración con herramientas analíticas. Estos hallazgos se alinean con lo señalado por Romero y Melendres (2023), quienes destacan que los Data Warehouses son fundamentales para decisiones basadas en evidencia y mejoran la eficiencia organizacional.

Sin embargo, se identificaron retos importantes:

- La necesidad de incorporar mecanismos avanzados de control de calidad de datos.
- La automatización de procesos ETL para reducir la intervención manual.
- La escalabilidad del sistema ante volúmenes crecientes de datos.

CONCLUSIONES

Se implemento el Data Warehouse para la evaluación de campos electromagnéticos en la ESPOCH demostrando ser una solución eficaz para integrar, almacenar y analizar grandes volúmenes de datos provenientes de mediciones temporales y espaciales. El modelo dimensional en estrella, junto con la infraestructura basada en PostgreSQL y PostGIS, permitió la optimización de consultas analíticas y facilitar la integración con herramientas de visualización como Power BI y ArcGIS, contribuyendo a la toma de decisiones estratégicas.

El empleo de Python para los procesos ETL constituyó una ventaja significativa frente a herramien-

tas tradicionales como Pentaho, tanto en términos de rendimiento como de flexibilidad. La facultad de personalizar transformaciones, incorporar cálculos específicos (factores de extrapolación y RMS) y automatizar validaciones básicas de calidad de datos confirma lo señalado por Kamlofsky et al. (2024), quienes destacan la eficiencia y simplicidad que ofrecen librerías como Pandas en tareas ETL.

Por otra parte, se evidenció la importancia de la calidad de datos como factor crítico para garantizar la confiabilidad del análisis. En este proyecto se implementaron controles básicos, futuras mejoras, que deben orientarse hacia la adopción de metodologías sistemáticas como TDQM y estándares internacionales como ISO/IEC 25012, tal como lo proponen Barreno & Silva (2025). Esto permitirá asegurar dimensiones clave como unicidad, completitud y conformidad, fortaleciendo la gobernanza de datos.

De manera concluyente, los resultados confirman la relevancia del Data Warehouse como una herramienta fundamental para la toma de decisiones basada en evidencia, alineándose con lo expuesto por Romero & Melendres (2023). La integración de datos heterogéneos, la generación de dashboards y mapas electromagnéticos, y la posibilidad de aplicar modelos predictivos aportan significativamente la gestión académica y técnica en la ESPOCH.

FINANCIACIÓN

La investigación fue financiada parcialmente por la Escuela Superior Politécnica de Chimborazo (ESPOCH) en el marco del proyecto EVALUACION DE LA EXPOSICION POBLACIONAL DE LOS CAMPOS ELECTROMAGNETICOS DE RADIO, TELEVISION Y CELULARES, PARA DESARROLLAR UNA HERRAMIENTA COMPUTACIONAL DE MAPEO ELECTROMAGNETICO EN LA ESPOCH IDIPI-284.

CONFLICTO DE INTERESES

Los autores declaran que no tienen conflictos de intereses relacionados con la presente investigación.

CONTRIBUCIÓN DE AUTORÍA

Rol	Danilo Llangarí	Pedro Infante
Participar activamente en:		
Conceptualización	X	X
Análisis formal	X	
Adquisición de fondos		
Investigación	X	X
Metodología	X	
Administración del proyecto	X	
Recursos	X	
Redacción –borrador original	X	

Redacción –revisión y edición	X	X
La discusión de los resultados	X	X
Revisión y aprobación de la versión final del trabajo.		X

REFERENCIAS

Almeida, E. C. de. (2011). Estudo de viabilidade de uma plataforma de baixo custo para data warehouse (Preprint). arXiv. https://doi.org/10.48550/arXiv.1108.0729

Barreno-Pilco, B. A., & Silva-Peñafiel, G. E. (2025). Calidad de datos en el Data Lake de una Fintech de Servicios Financieros, basado en las dimensiones de la norma ISO/IEC 25012. Código Científico Revista de Investigación, 6(1), 731–759. https://doi.org/10.55813/gaea/ccri/v6/n1/915

Dragomir, R. G. (2016). The role of the modern informatics systems in assisted decisions. Annals of Spiru Haret University. Economic Series, 14(4), 49–58. https://doi.org/10.26458/1445

Guerrero García, M. J., & Rodas Silva, J. (2022). Análisis comparativo de metodologías y herramientas tecnológicas para procesos de Business Intelligence orientado a la toma de decisiones. Informática y Sistemas: Revista de Tecnologías de la Informática y las Comunicaciones, 6(1), 38–44. https://doi.org/10.33936/isrtic.v6i1.4522

Inmon, W. H. (2005). Building the data warehouse (4.ª ed.). Wiley.

Jensen, S. K., Thomsen, C., Pedersen, T. B. & Andersen, O. (2021). pygrametl: A Powerful Programming Framework for Easy Creation and Testing of ETL Flows. En Transactions on Large-Scale Data- and Knowledge-Centered Systems XLVIII (pp. 45–84). Springer. https://doi.org/10.1007/978-3-662-63519-3_3

Kamlofsky, J., Manzano, F., & Lopez Yse, D. (2024). Tareas ETL más simples con Pandas: funciones útiles aplicadas sobre datos públicos. Revista Abierta de Informática Aplicada, 8(1), 47–70. https://doi.org/10.59471/raia2024204

Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., & Becker, B. (2008). The Data Warehouse Lifecycle Toolkit (2nd ed.). Wiley.

Kirmani, M. M. (2017). Dimensional Modeling Using Star Schema for Data Warehouse Creation. Oriental Journal of Computer Science and Technology, 10(4), 745–754. http://dx.doi.org/10.13005/ojcst/10.04.07

Romero-Chuquital, A., & Melendres-Velasco, J. J. (2023). Uso de Data Warehouse para la toma de decisiones empresariales: una revisión literaria. Revista Científica de Sistemas e Informática, 3(2), 1–23. https://doi.org/10.51252/rcsi.v3i2.543 (revistas.unsm.edu.pe)

Santoso, L. W., & Yulia. (2017). Data warehouse with big data technology for higher education. Procedia Computer Science, 124, 93–99. https://doi.org/10.1016/j.procs.2017.12.134

Silva-Peñafiel, G. E., Córdova-Vaca, A. M., Cusco-Vinueza, V. A., Cusco-Vinueza, V. A., & Estrada-Velasco, M. V. (2021). Implementación de un Data Warehouse mediante la metodología Hefestos para la toma de decisiones en el Instituto Nacional de Patrimonio Cultural Regional 3. Dominio de las Ciencias, 7(3), 1116–1135. https://doi.org/10.23857/dc.v7i3.2044

Souibgui, M., Atigui, F., Zammali, S., Cherfi, S., & Ben Yahia, S. (2019). Data quality in ETL process: A preliminary study. Procedia Computer Science, 159, 676–687. https://doi.org/10.1016/j.procs.2019.09.223

Vilela, F. de A., Times, V. C., Bernardi, A. C. de C., Freitas, A. de P., & Ciferri, R. R. (2023). A non-intrusive and reactive architecture to support real-time ETL processes in data warehousing environments. Heliyon, 9(5), e15728. https://doi.org/10.1016/j.heliyon.2023.e15728

Wijaya, W., Wiratama, J., & Wijaya, S. F. (2024). The implementation of data warehouse and star schema for optimizing property business decision making. G-Tech: Jurnal Teknologi Terapan, 8(2), 1242–1250. https://doi.org/10.33379/gtech.v8i2.4091

Ñauta Ñauta, F. O., & Silva Peñafiel, G. E. (2025). Data Warehouse para la toma de decisiones oportunas: caso de estudio Empresa de Agua Potable de Azogues (EMAPAL). Código Científico Revista de Investigación, 6(1), 710–730. https://doi.org/10.55813/gaea/ccri/v6/n1/914