



Escuela de Posgrados

**AUTOMATIZACIÓN DEL PROCESO DE ETL PARA DEFINIR EL
ACUERDO DE NIVEL DE SERVICIO MEDIANTE SQL SERVER
INTEGRATION SERVICES (SSIS)**

Bayron David Sánchez Muñoz

María Clara Martínez Vargas

Trabajo de Grado presentado como requisito para optar al título de:

Especialista en Big Data e Inteligencia de Negocios

Asesores:

Ingrid Durley Pardo Torres

Doctora en Ingeniería

Víctor Daniel Gil Vera

Doctor en Ingeniería de Sistemas

Universidad Católica Luis Amigó

Facultad de Ingenierías y Arquitectura

Especialización en Big Data e Inteligencia de Negocios

Medellín, Colombia

2024

Dedicatoria

A mis padres

María Clara Martínez Vargas

A mi familia

Bayron David Sánchez Muñoz

Agradecimientos

Agradecemos especialmente a nuestra compañera Catalina Ramos, quien estuvo presente desde el inicio de este desafiante camino, su aporte fue invaluable para nosotros.

También extendemos nuestro agradecimiento a la Universidad Católica Luis Amigó por proporcionar este valioso espacio de aprendizaje y conectarnos con profesionales que nos brindaron valiosas enseñanzas para este trabajo. En particular, queremos expresar nuestro reconocimiento al profesor Víctor Daniel Gil por su orientación inicial y a la profesora Ingrid Pardo por su acompañamiento que nos ayudó a culminar este proceso.

Finalmente, queremos expresar nuestro sincero agradecimiento a TIGO por brindarnos esta invaluable información, la cual ha sido fundamental para el desarrollo de nuestro proyecto de grado.

Resumen

Las empresas gestionan gran cantidad de datos de diversas fuentes para sus operaciones. Automatizar estas tareas es esencial para ahorrar tiempo y dinero, reducir errores y permitir a los empleados enfocarse en tareas prioritarias. Este proyecto se enfocó en implementar herramientas de Business Intelligence (BI) e Integration Services para automatizar procesos de integración de datos.

Para la empresa de telecomunicaciones, se subcontrató el servicio de soporte y desarrollo de informes para el área de BI y Big Data. La facturación mensual se basa en horas trabajadas. TIGO asigna tareas a la empresa subcontratada a través de una plataforma de tiquetes, donde se registran los tiempos trabajados y novedades. Estos datos son cruciales para calcular el cumplimiento del Acuerdo de Nivel de Servicio (ANS), que actualmente requiere el trabajo constante de tres personas, con una carga promedio de 4.5 horas diarias cada una.

Se implementó un proceso ETL en Microsoft SQL Server Integration Services (SSIS) para integrar datos de la plataforma de tiquetes y realizar cálculos mediante comandos SQL. Este proceso se automatizó mediante Jobs (tareas programadas) en la misma herramienta, garantizando la generación y envío del reporte final en intervalos determinados. Como resultado, se redujo el tiempo empleado en esta, de aproximadamente 4050 minutos a menos de 5 minutos por semana.

Palabras clave: Integration Services, Extracción Transformación y Carga (ETL), Automation, SQL Server Integration Services (SSIS), Acuerdos de Nivel de Servicio (ANS).

Nota aclaratoria: Los datos proporcionados han sido autorizados por la empresa TIGO exclusivamente para uso académico. Ninguna de las informaciones presentadas aquí representa la opinión o conocimiento directo de la empresa.

Tabla de contenido

1. Introducción	10
2. Planteamiento del problema	12
3. Justificación	14
4. Marco de referencias	17
5. Antecedentes	20
6. Objetivos	25
6.1 Objetivo general	25
6.2 Objetivos específicos	25
7. Viabilidad	26
8. Metodología	27
8.1 Etapa teórica	27
8.2 Etapa de implementación	28
9. Resultados	29
9.1 Objetivo 1: Identificación de datos y extracción de información	29
9.2 Objetivo 2: Transformación de datos	51
9.3 Objetivo 3: Carga de los datos e implementación del modelo	61
9.4 Objetivo general: Automatización en SSIS	64
9.5 Resultado adicional: Inteligencia de negocios	66
10. Conclusiones	69
11. Recomendaciones	70
12. Referencias	71

Lista de tablas

Tabla 1. <i>Comparativa artículos según sus resultados y objetivos</i>	23
Tabla 2. <i>Clasificación de las tareas y demás características</i>	29
Tabla 3. <i>Query para la extracción de datos hacia la tabla de cada tarea y responsable.</i> .	42
Tabla 4. <i>Query para identificar las OC únicas</i>	44
Tabla 5. <i>Query omisión de la carga de algunas OC mal tipificadas</i>	48
Tabla 6. <i>Query para la clasificación de tareas completadas y canceladas</i>	51
Tabla 7. <i>Query para el cálculo de horas trabajadas por responsable y categoría.</i>	52
Tabla 8. <i>Query Inserción de la fecha de creación y cierre de cada tiquete</i>	56
Tabla 9. <i>Query para inserción de las horas calculadas según la fecha de creación y cierre</i>	57
Tabla 10. <i>Query horas trabajadas totales</i>	57
Tabla 11. <i>Query para inserción de la prioridad e impacto para las mejoras continuas.</i>	58
Tabla 12. <i>Query para la verificación del cumplimiento del ANS.</i>	59
Tabla 13. <i>Query para la eliminación de OC mal tipificados.</i>	61

Lista de figuras

Figura 1. <i>Flujograma prisma.</i>	21
Figura 2. <i>Tiquetes con sus diferentes tareas.</i>	31
Figura 3. <i>Informe en el servicedesk.</i>	31
Figura 4. <i>Ingreso al escritorio remoto.</i>	32
Figura 5. <i>Tablas creadas en SSIS.</i>	33
Figura 6. <i>Conexión de la base de datos a SSIS.</i>	35
Figura 7. <i>Conexión vista desde los elementos del paquete SSIS.</i>	35
Figura 8. <i>Paquetes en SSIS para períodos determinados.</i>	35
Figura 9. <i>Ruta compartida en el servidor del DWH.</i>	36
Figura 10: <i>Herramientas del paquete SSIS.</i>	36
Figura 11: <i>Inicio flujo ETL</i>	37
Figura 12. <i>Conexión de cada elemento a la base de datos.</i>	37
Figura 13. <i>Formulación de la instrucción SQL en un elemento.</i>	38
Figura 14. <i>Carga del archivo .CSV a SSIS.</i>	38
Figura 15. <i>Carga de los datos a la tabla en SSIS.</i>	39
Figura 16. <i>Conexión y configuración de un elemento al archivo .CSV.</i>	39
Figura 17. <i>Conversión de las variables.</i>	40
Figura 18. <i>Proceso de mapeo.</i>	41
Figura 19. <i>Extracción de datos hacia la tabla de cada tarea y responsable.</i>	41
Figura 20. <i>Identificar las OC únicas.</i>	43
Figura 21. <i>Extracción de la información según el responsable y categoría.</i>	47
Figura 22. <i>Omisión de la carga de OC mal tipificadas.</i>	47
Figura 23. <i>Clasificación de tareas completadas y canceladas.</i>	51
Figura 24. <i>Cálculo de horas trabajadas por responsable y categoría.</i>	52
Figura 25. <i>Inserción de la fecha de creación y cierre de cada tiquete.</i>	56
Figura 26. <i>Inserción de las horas calculadas según la fecha de creación y cierre.</i>	56
Figura 27. <i>Horas trabajadas totales.</i>	57
Figura 28. <i>Inserción de la prioridad e impacto para las mejoras continuas.</i>	58
Figura 29. <i>Verificación del cumplimiento del ANS.</i>	59
Figura 30. <i>Eliminación de OC mal tipificados.</i>	61
Figura 31. <i>Creación de los archivos en Excel I.</i>	61
Figura 32. <i>Creación de los archivos en Excel II.</i>	62
Figura 33. <i>Archivos guardados en la ruta definida.</i>	62

Figura 34. <i>Elemento para el envío del correo electrónico</i>	63
Figura 35. <i>Configuración del elemento para el envío del correo</i>	63
Figura 36. <i>Correo final</i>	64
Figura 37. <i>Creación de Jobs</i>	64
Figura 38: <i>Configuración Job para el mes actual</i>	65
Figura 39: <i>Reprocesamiento para el mes actual</i>	65
Figura 40. <i>Configuración Job del mes anterior</i>	66
Figura 41. <i>Reprocesamiento para el mes actual</i>	66
Figura 42. <i>ANS fallas, incidentes y problemas</i>	67
Figura 43. <i>ANS mejoras continuas y solicitudes de servicio</i>	68

1. Introducción

Todas las empresas y organizaciones, independientemente de su sector de negocio, necesitan manejar y analizar una amplia variedad de datos provenientes de múltiples fuentes para llevar a cabo sus operaciones. Cuando se trata de grandes volúmenes de información, realizar estas tareas manualmente puede ser costoso en términos de tiempo y dinero, y puede conducir a posibles errores en el proceso. Además, muchas de estas tareas son repetitivas, lo que significa que se puede considerar la opción de automatizarlas para liberar a los empleados y permitirles concentrarse en otras áreas donde se pueda aprovechar mejor su talento (Seibold et al., 2022; Bunod et al., 2022).

Aunque, la automatización de la transformación de datos puede lograrse mediante un Extract, Transform and Load (ETL) creado con Integration Services, existen diversas estrategias de código abierto y soluciones comerciales disponibles para alcanzar el mismo objetivo. La elección de la mejor opción dependerá de las necesidades específicas, los recursos disponibles y el tamaño de la empresa o institución.

En primer lugar, se presentan estrategias de ETL para la mejora de la calidad de datos. Entre estas estrategias se encuentra Data Integration (TDI), una herramienta de ETL de código abierto. Por otro lado, Informatica Data Integration (IDI) y Microsoft SQL Server Integration Services (SSIS) son soluciones de ETL de código cerrado. Estas soluciones son capaces de filtrar registros duplicados mediante la transformación del agregador, la cual elimina duplicados utilizando agrupaciones de datos (Souibgui et al., 2019). Este enfoque resulta especialmente beneficioso para la limpieza de datos y para mejorar su calidad.

SSIS representa una solución empresarial de ETL que aborda desafíos complejos como la copia o descarga de archivos, la notificación por correo electrónico en respuesta a eventos, el análisis, almacenamiento y depuración de datos. Los paquetes de SSIS tienen la capacidad de operar de manera independiente o en colaboración con otros paquetes de SSIS. Esta herramienta puede extraer y transformar información de diversas fuentes, como archivos XML, CSV o bases de datos relacionales, y posteriormente cargarla en uno o múltiples destinos. Además, se distingue por ofrecer un entorno gráfico que simplifica la tarea, evitando la necesidad de escribir código complicado (Boulila et al., 2023). SSIS fue la herramienta seleccionada para este proyecto.

La migración, la recopilación y la automatización de datos son uno de los desafíos que pueden enfrentar las empresas al momento de empezar a incursionar en la mejora de sus procesos (Dogea & Stolt, 2021). Por lo tanto, en este trabajo se busca implementar una solución de automatización del proceso de ETL del Acuerdo de Nivel de Servicio (ANS). Esto, considerando las posibles complicaciones, mediante el análisis de proyectos afines y la revisión de la literatura pertinente para su ejecución, con el objetivo de atenuar obstáculos potenciales.

La automatización juega un papel fundamental en la obtención oportuna y precisa de los informes del ANS. En el pasado, se han registrado errores tanto intencionales como no intencionales al calcular estos cumplimientos, lo que consume un tiempo valioso para los miembros del equipo de trabajo de la empresa de telecomunicaciones y representa un costo económico para la organización. Esta automatización permitió que los informes se presenten con una mejor calidad a los responsables del proceso, para que así, puedan evaluar el desempeño en el cumplimiento de las tareas solicitadas al contratista.

Por ello, se implementó la herramienta SSIS para llevar a cabo todo el proceso de ETL. Este proceso comienza con la extracción de datos del documento CSV proporcionado por el subcontratista, seguido de la transformación de la información, donde se calculan las horas trabajadas tanto por el subcontratista como por la empresa de telecomunicaciones. Posteriormente, se evalúa el cumplimiento de los ANS y se genera un informe en formato Excel, el cual es enviado de manera automática a los interesados por correo electrónico. Todo este procedimiento se automatizó mediante los Jobs (tareas programadas) de la herramienta, configurando los tiempos de ejecución y envío para que se realice de manera periódica, ya sea mensual o semanalmente, según sea necesario.

2. Planteamiento del problema

UNE EPM Telecomunicaciones - TIGO es una entidad colombiana especializada en servicios de comunicaciones, respaldada por inversiones provenientes de dos importantes grupos, Millicom y EPM. A través de sus empresas afiliadas, que incluyen a Colombia Móvil, Edatel y Orbitel, Ofrece servicios de comunicación en áreas urbanas que abarcan a más del 70% de la población y tiene presencia en 767 municipios distribuidos en los 32 departamentos de Colombia (TIGO Colombia, 2023).

Desde la fusión con UNE, Edatel y ETP en 2014, la empresa ha seguido una estrategia de mantener su operación utilizando personal externo, para enfocarse en su actividad económica principal. Sin embargo, a lo largo de los años, esta decisión ha resultado en un aumento significativo en los costos operativos en todas las áreas del departamento de operaciones.

Por otro lado, dentro de las operaciones de la empresa, existe un equipo especializado en desarrollo y soporte encargado del Data Warehouse (DWH) y los ecosistemas de Big Data. Estos recursos se utilizan exclusivamente por las áreas legal, comercial y financiera de la compañía, permitiendo a estos departamentos acceder a información organizada de manera oportuna para tomar decisiones acertadas en asuntos externos a la organización.

Adicionalmente, dado que este equipo contaba únicamente con 10 especialistas para atender tanto el mercado móvil como el fijo. Por lo que, con el objetivo de garantizar el correcto funcionamiento de estos ecosistemas, la empresa contrató a un proveedor nuevo (empresa subcontratada) en el año 2016. Este subcontratista se encarga de brindar soporte y desarrollar todos los informes requeridos. La facturación se realiza mensualmente y está sujeta a los términos del contrato, que estipulan que el pago se basa en las horas de trabajo. La medida del rendimiento del contratista se basa en el tiempo dedicado a diversas tareas utilizando una herramienta de soporte y registro de información.

La empresa ha estado calculando estos costos de manera mayoritariamente manual, utilizando Microsoft Excel como herramienta para determinar cuántas horas se deben facturar cada mes. Además, se requiere la labor diaria de tres personas para llevar a cabo estos cálculos, lo que resulta en un promedio de 4.5 horas de trabajo al día por cada una

de ellas. Cuando se aborda la gestión de grandes volúmenes de datos, realizar estas tareas de forma manual puede resultar costoso en términos de recursos humanos y financieros, al tiempo que incrementa el riesgo de posibles errores en el proceso. Además, las tareas repetitivas impiden a los empleados concentrarse en otras áreas donde puedan aplicar sus habilidades de manera más eficaz (Seibold et al., 2022; Bunod et al., 2022).

En los últimos años, se ha observado un aumento significativo en la cantidad de horas trabajadas, no sólo por parte del subcontratista en el equipo de BI (Business Intelligence) y Big Data, sino también por otros contratistas en diferentes áreas de la operación. Este aumento se ha intensificado durante el año 2023. Hasta este momento, no se había considerado la posibilidad de aprovechar las herramientas de BI y Big Data para optimizar las operaciones internas de la empresa. Además, los desafíos que la empresa enfrenta relacionados con procesos de renovación del espectro radioeléctrico y ampliación de servicios han contribuido a que la empresa experimente un flujo de caja negativo. Esto ha llevado a que los accionistas tengan que inyectar capital por un monto cercano a los seiscientos mil millones de pesos colombianos para asegurar la continuidad de las operaciones de la compañía.

En respuesta a la situación anterior, la empresa ha emprendido campañas enfocadas en la optimización de recursos. En este contexto, se plantea la necesidad de implementar un control más eficiente de las horas trabajadas por los contratistas. Esto se debe a que, en la gestión de los tiquetes de trabajo, intervienen tanto los trabajadores externos como el personal interno de la compañía, y con el método actual de medición, no se puede garantizar al 100% la exactitud de las horas registradas por los contratistas, lo que podría resultar en pagos excesivos. Una gestión ineficiente de los datos puede constituir un obstáculo importante para una empresa. Por lo tanto, resulta esencial que los datos se reconozcan como un recurso valioso y se administren de manera acorde a su relevancia (Coito et al., 2020).

Dado esto, se ha identificado la necesidad de establecer un sistema de reportes detallado y preciso que pueda proporcionar información de manera más ágil. Esto permitirá tomar decisiones más informadas con respecto a la gestión de los contratistas dentro de la empresa, contribuyendo así a la reducción de costos.

3. Justificación

La automatización del proceso de integración de datos a través de SSIS y otras herramientas pueden presentar una serie de desafíos sustanciales para las organizaciones que buscan implementarla de forma efectiva. Entre los obstáculos más significativos se encuentra la complejidad de los datos.

Por ejemplo, en el caso del sector salud, se manipula una gran cantidad de datos provenientes de múltiples fuentes (hospitales, consultorios y clínicas en diferentes ubicaciones) que, a pesar de utilizar el mismo software para la administración de la información, se obtienen diferentes tipos de datos: Estructurados como códigos de diagnósticos y resultados de laboratorio, y no estructurados como las notas sobre los pacientes, escritas por los médicos y enfermeras. Por lo que, para implementar un sistema automatizado, primero se debe transformar los datos (estandarizarlos), lo cual implica la formulación de estrategias de comunicación y coordinación (Behnke et al., 2021).

La baja calidad de los datos tiene consecuencias tanto directas como indirectas en el proceso de toma de decisiones. Esto se origina debido a la falta de claridad en la información, conexiones con bases de datos desactualizadas y la constante evolución de las estructuras de las fuentes de datos a lo largo del tiempo (Souibgui et al., 2019).

No se trata de un desafío exclusivo para el sector de la salud, como señala Schmied et al (2020), el sector manufacturero también ha experimentado un crecimiento significativo, junto a una creciente diversidad en las fuentes de datos disponibles. En muchas ocasiones, la información se encuentra descentralizada (depende de una máquina, por ejemplo), lo que dificulta el acceso para otros participantes en el proceso. Por esta razón, resulta fundamental contar con un lenguaje estandarizado e implementar un sistema de información capaz de compartir datos y adaptarse a los cambios en los procesos de producción.

Para gestionar eficientemente el gran volumen y diversidad de datos disponibles en una organización, es posible emplear operaciones de ETL, mediante las cuales se revisa, depura y valida la información proporcionada. Sin embargo, la gestión de datos y el uso de estas herramientas requieren la experiencia de un especialista en el campo, lo que

puede dificultar su implementación en las empresas que carecen de dicha experiencia (Keskin & Yazıcı, 2023).

Por otro lado, tenemos las herramientas de BI, que juegan un rol importante en la integración de los datos. Pero, la complejidad y costos de los sistemas de BI dificultan su adopción en empresas, especialmente las medianas y pequeñas. Esto se debe a la necesidad de soluciones personalizadas, preocupaciones de seguridad y mantenimiento de herramientas (Kubina et al., 2015). Cuando una solución es excesivamente complicada, los usuarios pueden tener dificultades para comprenderla y utilizarla correctamente. Asimismo, si se requieren capacitaciones extensas o el proceso de aprendizaje es lento, existe un riesgo de que las instituciones pierdan su interés en su implementación (Hmoud et al., 2023).

Es importante tener en cuenta que la inversión en tecnologías de BI puede ser costosa, ya que implica gastos en infraestructura, software, licencias, formación y salarios. Además, es frecuente que muchas organizaciones no alcancen los beneficios esperados de BI (Gaardboe et al., 2017). Adicionalmente, los costos asociados con una solución de BI a lo largo del tiempo pueden aumentar debido a la necesidad de establecer estándares, aplicar gobernanza, salvaguardar los datos, garantizar la privacidad y mantener una mejora continua en las tecnologías utilizadas (Llave, 2017).

Como se ha evidenciado, los desafíos relacionados con la automatización no se ciñen a una herramienta específica ni a un sector empresarial en particular. Por lo tanto, resulta crucial examinar nuevas oportunidades de negocio y evaluar su impacto en las operaciones de la empresa. En este contexto, nos enfocaremos en el análisis de una empresa de telecomunicaciones y la automatización de su sistema de reporte de tareas mediante la herramienta SSIS.

La implementación de la automatización con SSIS se presenta como una solución viable, especialmente al considerar que los datos se originan en una base externa. El objetivo es asegurar la actualización constante de la información y evitar posibles inconvenientes relacionados con la calidad de los datos. Además, esta aproximación incluye la ejecución de los cálculos necesarios, como se detalla en el capítulo de resultados, lo que permite explorar nuevas capacidades de la herramienta SSIS.

Este proyecto abordó algunas de las problemáticas identificadas, ofreciendo un detallado paso a paso del proceso de desarrollo. Este enfoque resulta beneficioso para aquellas empresas que enfrentan desafíos similares y buscan una solución empleando la herramienta SSIS. La simplificación de este proceso busca mitigar problemas de sobrecostos o desistimiento debido a la complejidad del proyecto.

Además, la implementación de esta solución posibilitará a la empresa de telecomunicaciones reducir los tiempos asociados al control de las horas trabajadas, disminuyendo el esfuerzo dedicado por el personal a estas tareas y generando ahorros económicos. A su vez, se aborda el aspecto de la transparencia de datos, ya que la técnica propuesta agiliza el cálculo y la comparación de información, facilitando la identificación de su veracidad.

4. Marco de referencias

En esta sección, se abordarán los conceptos fundamentales que conforman el marco de referencia que orienta el trabajo. El propósito es proporcionar al lector una base informativa sólida sobre el proyecto y aclarar los términos pertinentes para facilitar una lectura más comprensible. Se plantean definiciones de temas técnicos clave, como metodologías en el ámbito de la minería de datos, software especializado y diversas técnicas empleadas

Minería de datos: Es el procedimiento para extraer y elegir datos con el fin de generar entendimiento y ofrecer a los usuarios una representación comprensible de estos, al interpretar la información, reunir y exhibir los patrones presentes en los datos.(Schuh et al., 2019).

Acuerdos de Nivel de Servicio (ANS): Son criterios, plazos y condiciones definidas para la provisión de productos o servicios de un proceso a usuarios internos y externos. Estos criterios se utilizan para evaluar la puntualidad del servicio (Función Pública - Gobierno de Colombia, n.d.).

Extract, Transform and Load (ETL): Software diseñado para recuperar datos de sistemas de origen, procesarlos y luego cargarlos en un repositorio de datos centralizado (Deist et al., 2020).

Data Warehouse (DW): Base de datos centralizada donde se unen datos de múltiples sistemas de origen dentro de una misma organización (Deist et al., 2020). Para TIGO, este lugar sirve como el punto central donde se recopila toda la información relevante de diversas fuentes y bases de datos. Posteriormente, esta información se procesa para crear nuevas bases de datos de acuerdo con los requisitos establecidos.

Imputación: Proceso para estimar o suponer valores faltantes en un conjunto de datos. Existen diferentes métodos para realizar una imputación de datos, que deben ser seleccionados según la situación y procedencia de los datos (Medina & Galván, 2007).

Inteligencia de Negocios (BI): Se refiere a una variedad de herramientas de análisis que brindan fácil acceso a información que respalda los procesos de toma de decisiones

(Machado et al., 2019). La inteligencia de negocios en TIGO está estrechamente vinculada con aspectos financieros, regulatorios y comerciales. Se emplea especialmente en la generación de informes finales, utilizando herramientas como Tableau y Power BI, con el objetivo de organizar y visualizar datos clave para la toma de decisiones. Este enfoque incluye la integración de información proveniente de diversas fuentes, con el fin de clarificar y presentarla de manera comprensible para el usuario final

Business Analytics (BA): Metodología de pensamiento estructurado que emplea herramientas y métodos tanto cualitativos como cuantitativos, junto con técnicas estadísticas computacionales, con el fin de examinar datos, adquirir percepciones, proporcionar información y facilitar la toma de decisiones (Power et al., 2018).

Structured Query Language (SQL): Lenguaje informático diseñado para la manipulación y gestión de conjuntos de datos y sus relaciones. Es ampliamente utilizado en sistemas de bases de datos relacionales como Microsoft Office Access. Se destaca como un estándar internacional respaldado por organizaciones de estándares reconocidas como ISO y ANSI (Microsoft, 2023).

SQL Server Integration Services (SSIS): Herramienta para desarrollar soluciones empresariales para abordar desafíos de integración y transformación de datos, incluyendo la copia, carga y limpieza de archivos y almacenamientos de datos, así como la gestión de objetos (Microsoft, 2021a). Para TIGO, esta herramienta es fundamental en todas las etapas de extracción, transformación y carga, siendo clave para todos los procesos del Data Warehouse.

Procesamiento Analítico en Línea (OLAP): Tecnología útil para organizar y analizar datos en empresas. En OLAP, los datos se dividen en cubos que son configurados por expertos en datos para facilitar la creación de informes y gráficos, lo que ayuda en la toma de decisiones empresariales estratégicas (Microsoft, 2021b)

Inteligencia Artificial (IA): Disciplina informática que se centra en el uso de computadoras y máquinas para reproducir las capacidades cognitivas y de toma de decisiones propias de los seres humanos, como la resolución de problemas (IBM, n.d.-c).

Data Quality: La calidad de los datos es crucial para las empresas, ya que los estándares de calidad de datos aseguran que las decisiones se basen en información precisa, lo que ayuda a alcanzar los objetivos comerciales (IBM, n.d.-a)

Tecnologías de la Información (TI): Conjunto de elementos y recursos empleados para adquirir, guardar, procesar, transferir y mostrar datos en formato digital. Esto abarca desde el hardware, como computadoras y servidores, hasta el software, las infraestructuras de comunicación y la administración de la información (IBM, n.d.-b).

5. Antecedentes

La Revisión Sistemática de Literatura (RSL) se realizó con el objetivo de presentar los avances en la investigación científica relacionados con la implementación de herramientas de BI e Integration Services para la combinación de datos y su conexión con la automatización de este proceso. En esta RSL, se llevó a cabo un estudio exhaustivo de la literatura científica disponible sobre herramientas de BI e Integration Services para la administración de datos. Se abordaron los principales temas e ideas relacionadas con la implementación de sistemas de automatización y de integración de información en diversos tipos de empresas, así como se exploraron las herramientas, ventajas y dificultades de dicha implementación. Además, esta revisión contribuyó significativamente a señalar el camino para investigaciones futuras en este campo, proporcionando una base sólida de conocimiento y una visión completa de los desarrollos previos en la integración de datos a través de estas herramientas.

La metodología RSL proporcionó un marco riguroso y estructurado para la recopilación, análisis y síntesis de la información relevante disponible en la literatura sobre un tema específico. Siguiendo los pasos delineados por Velásquez (2015), el proceso inició con la contextualización y justificación de la revisión, la formulación de preguntas de investigación y el diseño del protocolo de búsqueda. A continuación, se procedió con la fase de ejecución, donde se llevaron a cabo la búsqueda, selección, evaluación y síntesis de los artículos identificados. Finalmente, se culminó con la creación del presente capítulo, que consolidó toda la información necesaria.

Para la búsqueda, se tuvieron en cuenta tres preguntas de investigación, lo cual permitió enfocar la búsqueda en los temas que realmente tuvieran relevancia para este trabajo:

- ¿Cuáles estrategias de BI se han aplicado para la automatización de la integración de los datos?
- ¿Cuáles son los beneficios de automatizar los procesos de integración de datos utilizando la herramienta Integration Services de BI en el ámbito empresarial?
- ¿Cuáles son los principales obstáculos que las organizaciones encuentran al intentar automatizar sus procesos de integración de datos y/o al implementar estrategias de BI?

Por otro lado, se utilizó el Reporte Preferidos para Revisiones Sistemáticas y Meta-Análisis (PRISMA – por sus siglas en inglés) fue desarrollado con el propósito de ayudar a los investigadores a presentar de forma transparente el motivo de la revisión, las acciones llevadas a cabo por los autores y los hallazgos obtenidos. Este, ofrece una orientación para la presentación de informes sobre revisiones sistemáticas que define los métodos para identificar, seleccionar, evaluar y sintetizar estudios. El diagrama de flujo representa visualmente el proceso de información a través de las distintas fases de una revisión sistemática. Detalla la cantidad de registros identificados, incluidos y excluidos, así como las razones de las exclusiones (Page et al., 2021).

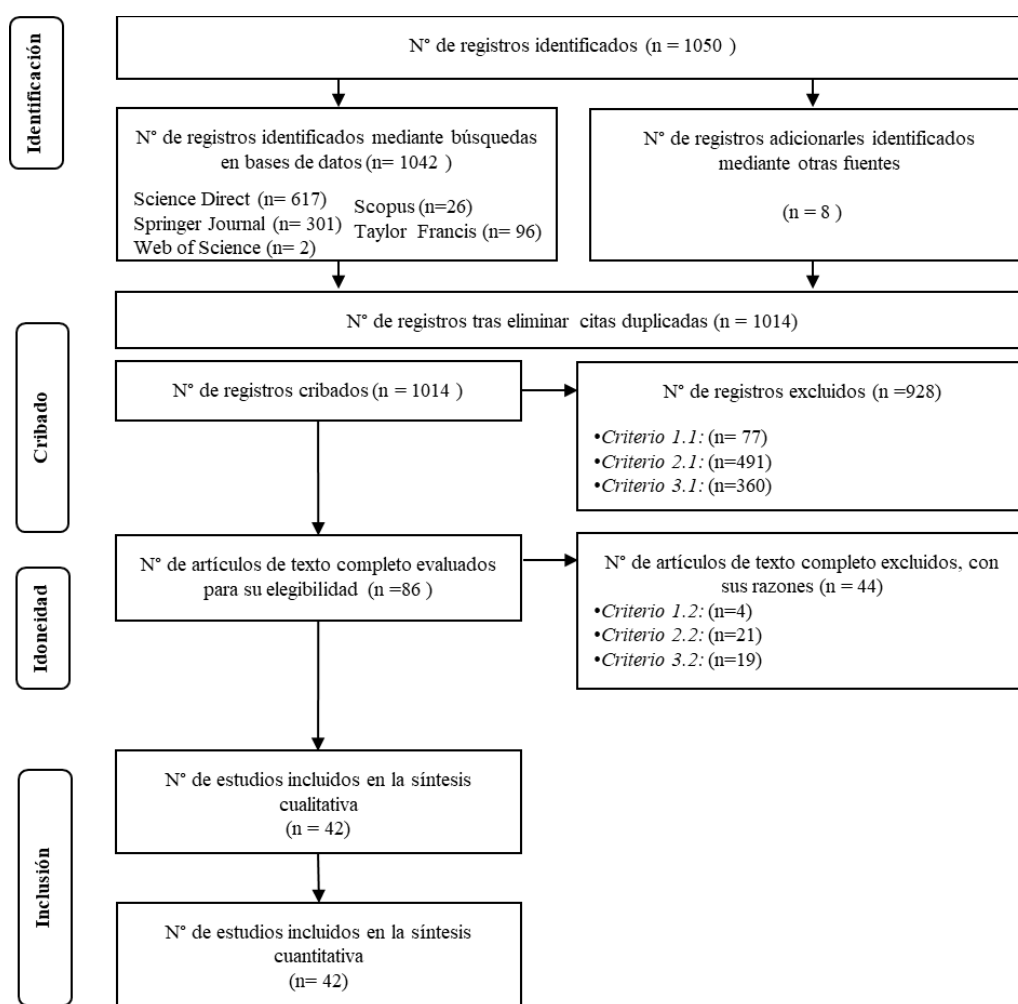


Figura 1. Flujograma prisma.

Por medio de Flujograma Primas (Figura 1), se pudo realizar una selección más adecuada de los documentos a estudiar. Se identificaron 42 artículos y otras fuentes pertinentes con la presente investigación y se revisaron en detalle, con el propósito de profundizar en el

estado actual de la automatización de procesos ETL y su optimización a través de herramientas de BI. A pesar de abordar el tema desde diversas perspectivas en el ámbito empresarial, en general se obtuvo un resultado consistente al analizar las motivaciones y objetivos detrás de la implementación de este tipo de proyectos.

Uno de los artículos con mayor similitud al de este trabajo fue realizado por Wu et al (2022), en el cual se muestra las etapas para la creación de un dashboard que integra los datos proporcionados por un departamento de otorrinolaringología pediátrica. El propósito principal fue agilizar la recopilación de información, reducir costos, y mejorar el acceso y la legibilidad de los datos. Esto, a su vez, fomentó la eficaz implementación de iniciativas destinadas a mejorar la calidad y la investigación en los servicios. A través de la utilización de la herramienta SSIS, se automatizó la recopilación y combinación de datos de forma semanal, los cuales nutren el mencionado dashboard.

Para la construcción de ese prototipo para la integración automática de registros a la base de datos, se requirió la participación de un experto para el proceso inicial y análisis posteriores, lo cual fue uno de los componentes claves de la inversión del proyecto. Además, otro de los costos se relaciona con el tiempo empleado, ya que el desarrollo se extendió durante nueve meses. Por otro lado, se incluyó el costo del ancho de banda debido a la gran cantidad de procesamiento de datos necesario (Wu et al., 2022). Sin embargo, para nuestro proyecto, no se realizará el cálculo de los costos asociados para el desarrollo de esta implementación.

Se tiene un estudio similar presentado por Deist et al (2020), en el que se contaba con la disponibilidad de varias bases de datos que abarcan a más de 20 000 pacientes distribuidos en 5 países y se requería unificar esta información. Estos pacientes eran gestionados en función de la etapa evolutiva de su condición, pero, en ocasiones, no se disponía de la información completa. Para abordar este desafío, se lleva a cabo la imputación de datos, que en este caso implica inferir lógicamente el valor según la información de las etapas anteriores y posteriores. Además, era necesario llevar a cabo análisis estadísticos resumidos a través de todas las bases de datos, con el fin de anticipar la supervivencia de los pacientes después del tratamiento. Este proceso puede compararse con el propósito de nuestro proyecto, ya que implica realizar cálculos previos a la consolidación de la información.

También, se tiene el caso del uso de la automatización para el control de carga que presenta Václav et al (2021) para generar informes en línea. En estos, se procesan los datos intermodales sobre trenes y otros medios de transporte, revisando la carga en unidades de contenedores y trenes. Además, utilizan las estrategias de BI para la toma de decisiones y la planificación empresarial. Esta metodología aumenta la eficiencia, identifica oportunidades de mejora, advierte sobre costos adicionales y puede ser un factor clave para el éxito en entornos competitivos como lo es el transporte de mercancía.

El presente proyecto comparte su enfoque con los ejemplos anteriores en el tema automatización, al mismo tiempo que prioriza la reducción del tiempo invertido y la mitigación de posibles errores inherentes a procesos manuales. Pero, adicionalmente, se concentra en la transparencia en los datos. Esto es esencial para el éxito financiero de esta empresa, puesto que esto incide directamente en los costos, específicamente de las horas trabajadas y las tareas completadas. Asimismo, se considera el factor humano, no sólo en términos de error, sino también en cuanto a la posible manipulación intencionada de datos por parte de los contratistas en su propio beneficio.

Por otro lado, aunque las herramientas de BI son valiosas para el análisis en diversos contextos, es importante destacar que su principal utilidad no suele ser la evaluación del rendimiento de los recursos humanos, caso contrario este trabajo. En su lugar, se centran en facilitar la toma de decisiones y proporcionar un acceso más ágil a la información. En la Tabla 1, se presenta un resumen del comparativo entre los estudios mencionados.

Tabla 1. Comparativa artículos según sus resultados y objetivos

	Automatización integración datos	BI para la toma de decisiones	Reducción tiempo	Mitigación errores	Transparencia en los datos	Evaluación rendimiento RRHH
Utility of the finance-electronic medical record digital dashboard in pediatric otolaryngology Wu et al (2022)	✓	✓	✓	✓	X	X
Distributed learning on 20 000+ lung cancer patients – The Personal Health Train Deist et al (2020)	✓	✓	✓	✓	X	X
Utilization of business intelligence tools in cargo control Václav et al	✓	✓	✓	✓	X	X

(2021)						
Automatización del Proceso ETL del Acuerdo de Nivel de Servicio mediante SSIS	✓	✓	✓	✓	✓	✓

De esta manera, se formuló la pregunta principal de investigación para el proyecto:
¿Cómo automatizar el proceso de entrega de informes del ANS relacionados con el desempeño en la subcontratación de servicios de soporte en la empresa TIGO?

6. Objetivos

6.1 Objetivo general

Automatizar el proceso de entrega de informes del Acuerdo de Nivel de Servicio (ANS) relacionados con el desempeño en la subcontratación de servicios de soporte en una empresa de telecomunicaciones utilizando SQL Server Integration Services (SSIS).

6.2 Objetivos específicos

- Extraer la información del trabajo realizado por el subcontratista del sistema de tiquetes de gestión “servicedesk” de la empresa
- Transformar los datos recopilados de acuerdo con los tipos de ANS.
- Cargar la información requerida para el envío por correo electrónico de los informes del ANS al área responsable.

7. Viabilidad

Este proyecto se enfocó en evaluar y dar seguimiento al contrato con subcontratista utilizando el análisis de los datos con herramientas de BI. Los recursos necesarios para llevar a cabo este proyecto incluyeron un equipo con experiencia en desarrollo y análisis de información, así como licenciamientos de SQL Server y Microsoft Windows Server, DWH, un equipo de cómputo con un procesador de 8 núcleos y 16 GB de RAM, análisis de información, solicitud de conexiones entre sistemas, gestión de permisos para descargar información y una conexión a internet sólida de mínimo 100 Megas para garantizar un funcionamiento eficiente.

El alcance del proyecto implicó determinar si el subcontratista cumplía con las horas contratadas para brindar soporte en el departamento de BI y Big Data. Esto involucró el desarrollo de dos ETL en IS, uno para controlar los cumplimientos del mes en curso y otro para el mes anterior. También, se necesitaron accesos a la plataforma de tiquetes que el contratista utiliza para dar evidencias del cumplimiento de su trabajo. Al contratista se le asignan tareas de desarrollo como paquetes del PL SQL, ETL, cubos, dashboard, reportes en archivos planos, DAG para Big data e integraciones en AWS. El presente proyecto abarcó un período de seguimiento de aproximadamente dos meses y la revisión del historial de información generada por el subcontratista.

Sus implicaciones incluyeron la automatización de procesos que solían ser manuales, lo que aumentó la eficiencia y reducción de los costos a largo plazo. También, al mantener un registro detallado del tiempo y las actividades del subcontratista, se pudo aumentar la transparencia y el control.

La implementación de este proyecto potencializa soluciones positivas para la empresa, principalmente relacionadas con el trabajo en conjunto con los contratistas. Se tiene un mayor control sobre sus entregables, siendo una motivación para el aumento en su desempeño tanto en la calidad como en la rapidez de su trabajo. También, al automatizar los procesos, se redujo la dependencia de operarios para alimentar y calcular información, liberando a los individuos para que realicen otras tareas, lo cual también implica una mejor económica. Además, al entregar informes de manera regular y programada, se pudo mejorar la eficiencia y se contribuye a los objetivos de la empresa. La seguridad de los datos también aumentó al reducir su manipulación por parte de los operarios.

8. Metodología

Este capítulo documenta el paso a paso realizado para dar cumplimiento a los objetivos específicos. La investigación se realizó en dos etapas: La etapa macro, que consiste en la exploración teórica de los conceptos y antecedentes relacionados con la temática del trabajo, y la etapa micro, que consiste en el proceso realizado como equipo, en el cual se utilizaron técnicas para la extracción, transformación, carga de datos, para la implementación final de la solución al problema planteado.

8.1 Etapa teórica

El primer paso en la ejecución de este proyecto consistió en investigar trabajos previos relacionados con la temática. Fue fundamental examinar procesos similares para determinar la viabilidad de esta implementación, empleando herramientas análogas a las establecidas en el presente trabajo. Adicionalmente, uno de los propósitos de este proyecto era desarrollar una solución que agregara valor a la empresa, por lo que el análisis de implementaciones similares en diversas industrias fue uno de los enfoques de la investigación

Para obtener esta información, se definieron palabras clave relevantes, como "Automatización" y "SQL Server Integration Services", y se realizaron búsquedas en bases de datos electrónicas como Scopus o Science Direct. Estas búsquedas se llevaron a cabo en inglés para ampliar nuestro alcance al buscar en las fuentes de información.

Luego, se definieron cadenas de búsqueda y ecuaciones utilizando conectores booleanos y filtros para la antigüedad de los artículos, garantizando así la obtención de los resultados más relevantes. Para evaluar estos resultados, se utilizó la herramienta visual flujograma Prisma (Figura 1), que permitió identificar de manera adecuada los artículos pertinentes.

Finalmente, se elaboró una tabla comparativa (Tabla 1) con los cinco artículos que presentaban la mayor similitud con el propósito de este trabajo. En esta tabla, se examinaron los puntos en común y se identificaron las áreas en las que el proyecto ofrece una contribución notable.

8.2 Etapa de implementación

Después de investigar sobre los conceptos y casos reales relacionados con la automatización e integración de datos, se procedió con el desarrollo de la aplicación. En la primera fase, se extrajo la información del sistema de tiquetes de gestión de la empresa, operado por el subcontratista, conocido como servicedesk. Este sistema contiene todos los datos necesarios para calcular los ANS, incluyendo información sobre tareas, responsables, y tiempos. La información se encontraba en un archivo .CSV, y se utilizó la herramienta SSIS para extraerla y organizarla en tablas.

El siguiente paso consistió en integrar y cargar la información obtenida a un archivo en Excel, ya que el entregable final de este proceso era generar un documento que permitiera a los encargados del área de la empresa de telecomunicaciones evaluar el desempeño del subcontratista.

Posteriormente, se transformaron los datos recopilados para calcular los tiempos requeridos para resolver un pedido, diferenciando entre tiempos utilizados por el subcontratista y tiempos de impedimento por parte de la empresa de telecomunicaciones. Los resultados de estos cálculos, comparados con los límites establecidos en la Tabla 2 permitieron clasificar si se cumplió o no con los estándares, lo que facilitó la evaluación del desempeño del subcontratista.

El proceso descrito anteriormente se realizó manualmente. Sin embargo, el principal objetivo de este proyecto era implementar un modelo automatizado de ETL a través de SSIS. Para lograrlo, se desarrollaron unos Jobs que ejecutan este proceso de forma programada, incluyendo la generación del reporte de los ANS en fechas específicas del mes y su envío por correo electrónico a los interesados. Con la información obtenida de estos reportes, se crearon dos gráficas en Power BI (Figura 43, Figura 42) para analizar el desempeño de cada tarea a lo largo del tiempo.

9. Resultados

En este capítulo, se detalla el desarrollo de la aplicación utilizada para alcanzar los objetivos específicos establecidos. Comienza con la fase de extracción de datos, la cual abarca todo el proceso de preprocesamiento de la información. Aquí se proporciona una explicación detallada sobre el origen, las características y forma de extracción de los datos iniciales empleados en la construcción del modelo. Posteriormente, se aborda la etapa de transformación de datos, donde se describe el proceso de eliminación, transformación o importación de variables específicas. Finalmente, se lleva a cabo la implementación automatizada de los procesos mencionados anteriormente, junto con la fase de carga de datos, para así, generar el informe final requerido por los usuarios, utilizando la herramienta.

9.1 Objetivo 1: Identificación de datos y extracción de información

La empresa de telecomunicaciones subcontrató servicios de soporte y desarrollo de informes para su área de BI y Big Data. La facturación mensual se determina según las horas trabajadas. TIGO asigna tareas a la empresa, las cuales siguen las características definidas en la Tabla 2.

Tabla 2. Clasificación de las tareas y demás características.

Categoría	Tipo	Tiempo	Detalles (tiempo hábil para dar respuesta)
7 x 24	Fallas	9 horas	Jornada continua – 24 horas
Correctivo programado	Problemas	27 horas	Jornada laboral – 7:00 am a 5:30 pm de lunes a viernes
	Fallas	9 horas	
	Incidentes	45 horas	
Configuraciones sistemas corporativos	Solicitudes	18 horas	
Plan mejora continua de aplicaciones	Complejidad Alta	12 días	
	Complejidad Media	9 días	
	Complejidad Baja	7 días	

La información relacionada con las tareas resulta crucial para calcular el cumplimiento del ANS, dado que los límites de tiempo para su finalización están claramente definidos. Para garantizar un seguimiento adecuado del cumplimiento de las tareas dentro de estos límites, se emplea un sistema de tiquetes denominado "servicedesk".

Cada ticket creado conlleva múltiples tareas, configuradas según las categorías establecidas previamente (ver Tabla 2). La configuración puede comprender una o varias tareas, asignadas ya sea a tiempo completo o parcial, junto con la identificación de los responsables de cada tarea. En el servicedesk, se registra la fecha de creación del ticket como la fecha de asignación de la primera tarea. Las tareas se llevan a cabo de manera secuencial; cuando el proveedor completa la primera, automáticamente se registra la fecha de cierre y se inicia la tarea siguiente, indicando su fecha de apertura.

Asimismo, en el servicedesk se documentan los tiempos de impedimento, los cuales definen el período durante el cual el proveedor debe esperar a que TIGO proporcione respuestas o cumpla con su parte correspondiente. Una vez recibidas las respuestas, el proveedor puede proceder con las tareas subsiguientes.

Es importante destacar que este tiempo de impedimento no se considera al evaluar el desempeño del proveedor respecto al cumplimiento de la tarea. Si una tarea no es responsabilidad del subcontratista, el proceso será igualmente aplicable al responsable designado de dicha tarea, en este caso, TIGO. La Figura 2 ilustra un ejemplo de un ticket (número 20007654) junto con sus respectivas tareas, comenzando desde la secuencia 10, incluyendo su descripción y asignación correspondiente.

En la Figura 3 se presenta el programa que facilita el reporte de información del servicedesk, permitiendo la descarga de datos de todos los tickets atendidos durante un período específico. Estos datos se exportan inicialmente en un archivo Excel y posteriormente se convierten en formato CSV, el cual se almacena en una ruta compartida dentro del DWH. Este conjunto de datos constituye la base sobre la cual se construye el desarrollo de este proyecto.

Seq	Task	Description	Assignee	Group	Status	Start Date	Completion Date
10	Validar Solicitud	El especialista técnico debe validar que la solicitud esté completa, sea clara y aplique como una solicitud de soporte y no del ciclo de vida de desarrollo de TI ni Desarrollos menores. Antes de completar esta tarea debe estar en su versión final el blueprint y adjunto.		Grupo ET - BI BSD	Complete	05/26/2021 09:49 am	05/26/2021 09:52 am
15	Group Start Task				Complete	05/26/2021 09:52 am	05/26/2021 09:52 am
16	Aval Líder Funcional SOX	Validar el listado de aplicaciones y registrar el aval funcional		Grupo líderes funcionales BI	Approve	05/26/2021 09:52 am	05/26/2021 11:09 am
17	Aval Líder Técnico SOX	Validar el listado de aplicaciones y asignar el aval líder técnico de acuerdo al listado de aplicaciones		Grupo líderes técnicos BI	Approve	05/26/2021 09:52 am	05/26/2021 10:15 am
18	Group End Task				Complete	05/26/2021 11:09 am	05/26/2021 11:09 am
20	Realizar documento de entendimiento	El proveedor debe realizar el documento de entendimiento y lo anexa a la OC.		Grupo Proveedor Soporte BI BSD	Complete	05/26/2021 11:09 am	05/26/2021 12:09 pm
30	Aprobar documento de entendimiento	El especialista técnico debe aprobar el documento de entendimiento generado por la fabrica.		Grupo ET - BI BSD	Approve	05/26/2021 12:09 pm	05/26/2021 02:52 pm
40	Elaborar Estimación	El Proveedor debe definir la estimación para la atención de la solicitud y antes de completar la tarea debe incluir el valor de la estimación en las propiedades de la OC.		Grupo Proveedor Soporte BI BSD	Complete	05/26/2021 12:14 pm	05/26/2021 02:52 pm
50	Aprobar Estimación	El especialista técnico debe aprobar la estimación generada por la fabrica.		Grupo ET - BI BSD	Complete	05/26/2021 02:52 pm	05/26/2021 02:52 pm
60	Informar Fecha Compromiso	El Proveedor debe definir la fecha de compromiso de entrega de la solución y antes de completar la tarea debe incluir la fecha de compromiso de entrega de la Solución en las propiedades de la OC.		Grupo Proveedor Soporte BI BSD	Complete	05/26/2021 02:52 pm	05/27/2021 02:57 pm
70	Elaborar Documentos de Análisis y Diseño	El Proveedor debe elaborar el documento de análisis y diseño.		Grupo Proveedor Soporte BI BSD	Complete	05/26/2021 02:57 pm	05/27/2021 08:06 am
80	Aprobar Análisis y Diseño Especialista Técnico	El Especialista técnico debe aprobar el documento de análisis y diseño. Si el proveedor no lo adjuntó porque no se requería, debe darle skip a la tarea o rechazarla para que el proveedor genere el entregable.		Grupo ET - BI BSD	Approve	05/27/2021 08:06 am	05/27/2021 10:04 am
90	Aprobar Análisis y Diseño Solicitante	El Solicitante debe aprobar el documento de análisis y diseño.	Rios Solano, Johana		Complete	05/27/2021 10:04 am	05/27/2021 10:04 am
100	Desarrollar Solución y Ejecutar Pruebas	El Proveedor debe desarrollar la solución y ejecutar las pruebas correspondientes y debe adjuntar las evidencias de las pruebas y manual de usuario cuando se requiere		Grupo Proveedor Soporte BI BSD	Complete	05/27/2021 10:04 am	05/27/2021 10:19 am

Figura 2. Tiquetes con sus diferentes tareas.

No es posible establecer una conexión directa con la base de datos del servicedesk debido a que se trata de una aplicación transaccional, la cual está diseñada para ser de uso general dentro de la compañía. Se ha adoptado la medida de evitar que aplicaciones externas realicen consultas directas a la base de datos, ya que consultas extensas o mal diseñadas podrían afectar el funcionamiento de la aplicación, impactando a todas las áreas que dependen del servicedesk. Por esta razón, es imperativo utilizar el generador de reportes (Figura 3) provisto por esta herramienta, el cual cuenta con consultas predefinidas.

Informe Tareas_OC Correctivos Cerrados

Total de tareas : 1,834

Orden de cambio	Fecha de Inicio OC	Change Close Date	Impact Symbol	Change Priority Symbol	Estado de OC	Sequence	Tarea	Task Description	Fecha Inicio Tarea	Fecha Final Tarea	Estado tarea
20915462	11/14/2023 11:12:56 AM	01/26/2024 04:32:58 PM	3 Medio	3 Media		2,375	Impedimento		12/22/2023 09:38:09 AM	01/05/2024 11:05:33 AM	COMP
20915462	11/14/2023 11:12:56 AM	01/26/2024 04:32:58 PM	3 Medio	3 Media		2,450	Cierre-Validacion de la		01/05/2024 11:05:33 AM	01/09/2024 04:07:37 PM	COMP
20915462	11/14/2023 11:12:56 AM	01/26/2024 04:32:58 PM	3 Medio	3 Media		2,525	Impedimento		01/09/2024 04:07:37 PM	01/16/2024 11:58:15 AM	COMP
20915462	11/14/2023 11:12:56 AM	01/26/2024 04:32:58 PM	3 Medio	3 Media		2,600	Post Mortem		01/16/2024 11:58:15 AM	01/22/2024 04:08:57 PM	COMP
20915462	11/14/2023 11:12:56 AM	01/26/2024 04:32:58 PM	3 Medio	3 Media		2,675	Impedimento		01/22/2024 04:08:57 PM	01/23/2024 06:44:41 PM	COMP

Figura 3. Informe en el servicedesk.

A continuación, se presentan la explicación de los atributos:

- Orden de cambio (OC): Identificador del tickete.
- Fecha inicio OC: Fecha cuando se crea el tickete.

- Change Close OC (Fecha de cierre OC): Fecha cuando se finaliza el ticket.
- Impact Symbol (Nivel de impacto): Es Alto, Medio o Bajo. Indica la afectación que tiene el cumplimiento de la petición en las operaciones de la empresa.
- Change Priority Symbol (Nivel de prioridad): Define la prontitud con la que se debe entregar la solución de la petición (Muy bajo, bajo, medio, alto).
- Sequence (Secuencia): Identificador de la tarea.
- Tarea: Nombre de la tarea. Se pueden tener tareas nombradas post mortem, que fueron creadas manualmente, la cual es usada para documentar un paso producción fallido. Se encuentra en la categoría de Impedimento.
- Fecha inicio tarea: Fecha de creación de la tarea.
- Fecha cierre tarea: Fecha de finalización de la tarea.
- Estado de la tarea: Puede ser – completado, cancelado, aprobado, rechazado, skipped.
- Minutos tiempo hábil: Cantidad de minutos hábiles en que la tarea estuvo abierta y cerrada.
- Change category symbol (Categoría): Nombre de la categoría (Definidas en la Tabla 2).

Con la información del reporte de servicedesk, se comenzó con el proceso de extracción de los datos, pero desde el desarrollo en SSIS.

Para ingresar al servidor de desarrollo del DWH, se hace a través de escritorio remoto (Figura 4). Este modo de ingreso es definido por políticas de la empresa.

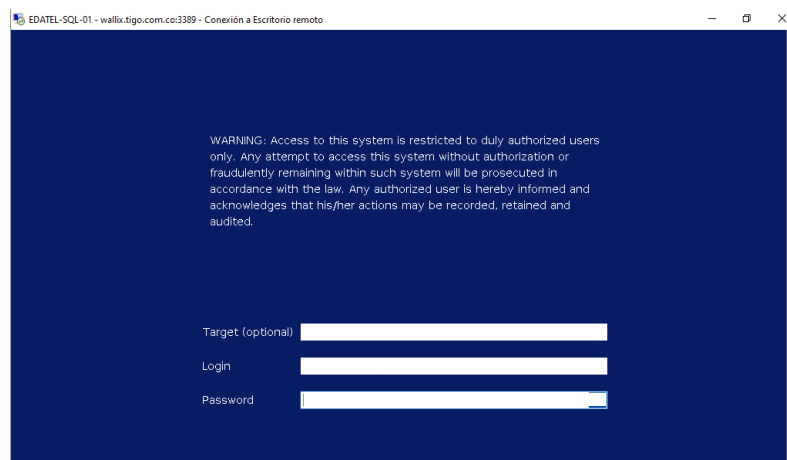


Figura 4. Ingreso al escritorio remoto.

Se creó la base de datos ANST en el DWH para extraer la información necesaria de la tabla original con el ETL. En la Figura 5 se muestran las nuevas tablas creadas.

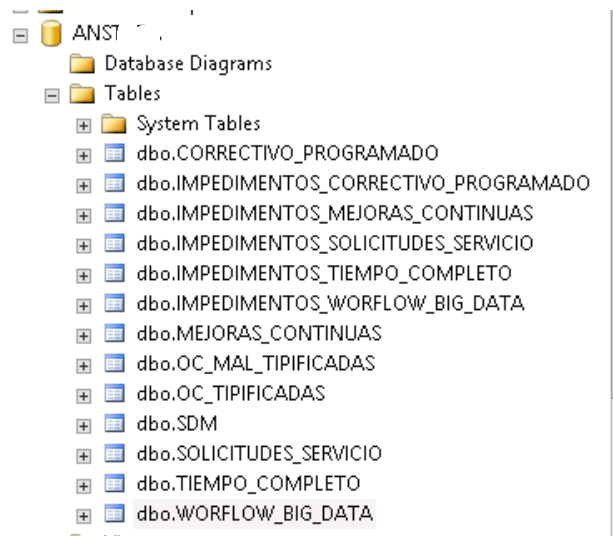


Figura 5. *Tablas creadas en SSIS.*

- SDM: Es la tabla de partida que almacena la información del dataset que luego será tomada por el flujo SSIS.
- CORRECTIVO_PROGRAMADO: Almacena la información de las tareas trabajadas por el proveedor de los tiquetes correctivos programados. Esta será tomada desde la tabla SDM.
- IMPEDIMENTOS_CORRECTIVOS_PROGRAMADOS: Almacena la información de las tareas trabajadas por TIGO de los tiquetes correctivos programados. Esta será tomada desde la tabla SDM.
- WORKFLOW_BIG_DATA: Almacena la información de las tareas trabajadas por el proveedor de los tiquetes correctivos programados para los desarrollos de Big data que se realizan dentro de la empresa. Esta será tomada desde la tabla SDM.
- IMPEDIMENTOS_ WORKFLOW_BIG_DATA: Almacena la información de las tareas trabajadas por TIGO de los tiquetes correctivos programados para los desarrollos de Big data que se realizan dentro de la empresa. Esta será tomada desde la tabla SDM.
- MEJORAS_CONTINUAS: Almacena la información de las tareas trabajadas por el proveedor de los tiquetes de mejoras continuas. Esta será tomada desde la tabla SDM.
- IMPEDIMENTOS_MEJORAS_CONTINUAS: Almacena la información de las

tareas trabajadas por TIGO de los tickets de mejoras continuas. Esta será tomada desde la tabla SDM.

- SOLICITUDES_SERVICIO: Almacena la información de las tareas trabajadas por el proveedor de los tickets de solicitudes de servicio. Esta será tomada desde la tabla SDM.
- IMPEDIMENTOS_SOLICITUDES_SERVICIO: Almacena la información de las tareas trabajadas por TIGO de los tickets de solicitudes de servicio. Esta será tomada desde la tabla SDM.
- TIEMPO_COMPLETO: Almacena la información de las tareas trabajadas por el proveedor de los tickets de 7 x 24. Esta será tomada desde la tabla SDM.
- IMPEDIMENTOS_TIEMPO_COMPLETO: Almacena la información de las tareas trabajadas por TIGO de los tickets de 7 x 24. Esta será tomada desde la tabla SDM.
- OC_MAL_TIPIFICADAS: Almacena la información de los tickets que, desde un inicio, no fueron creados correctamente.
- OC_TIPIFICADAS: Almacena la información resultante para ser entregada a los líderes e interventores. Los datos de esta tabla son resultados de las transformaciones que se hacen con la información de las tablas anteriores.

Se realizó la conexión desde el paquete SSIS a la base de datos ANST de SQL Server. De esta manera, se accedió a las tablas anteriormente mencionadas desde el paquete de SSIS (Figura 6).

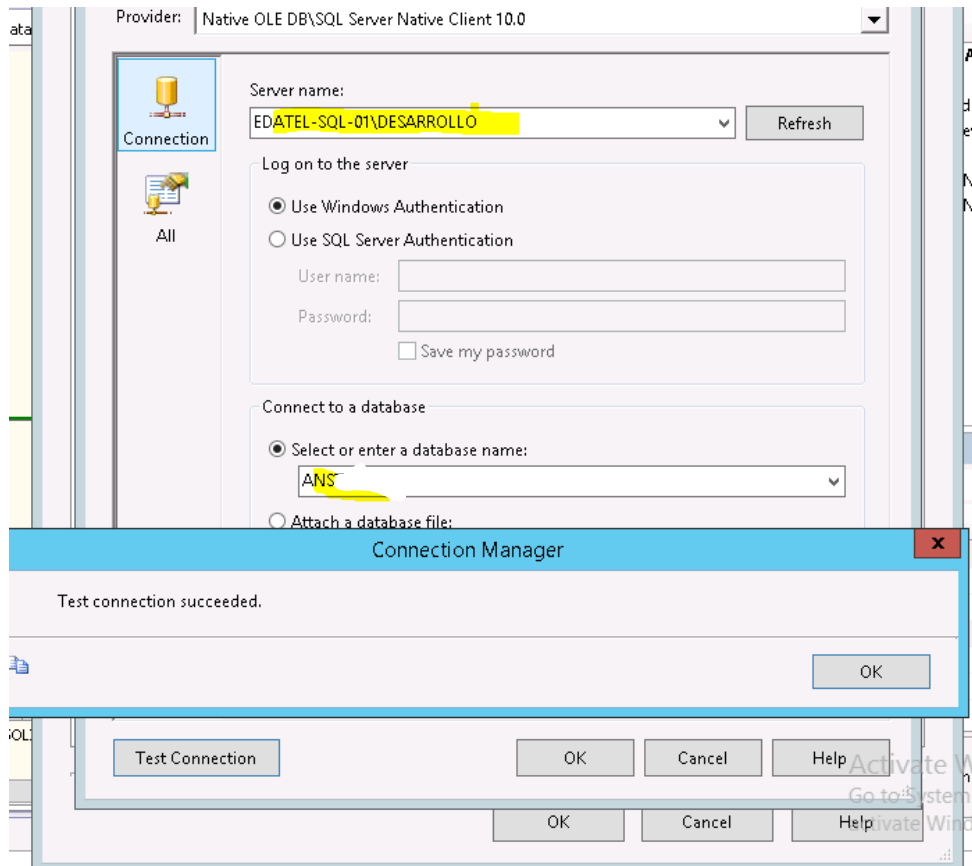


Figura 6. Conexión de la base de datos a SSIS.

La conexión vista desde los elementos del paquete SSIS (Figura 7).



Figura 7. Conexión vista desde los elementos del paquete SSIS.

Se crearon dos paquetes, uno para obtener la información cuando el mes finaliza y otro para obtener la información semanal del mes actual (Figura 8).



Figura 8. Paquetes en SSIS para períodos determinados.

Los dos dataset (uno para el mes actual y otro para el mes anterior), se guardan todos los viernes en una ruta compartida en el servidor del DWH, para utilizarlo en la obtención de los informes (ver Figura 9).

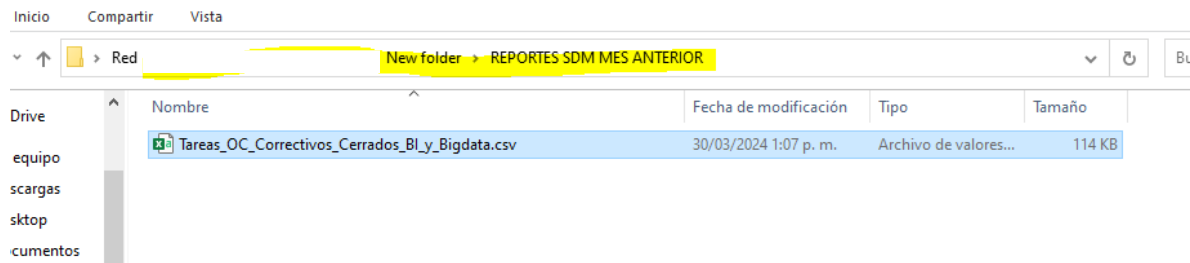


Figura 9. Ruta compartida en el servidor del DWH.

El siguiente paso consistió en la creación del proceso ETL utilizando el paquete SSIS. Para llevar a cabo esta tarea, se emplearon diversas herramientas (ver Figura 10) que facilitaron la conexión con las tablas de la base de datos, así como la realización de consultas y transformaciones de la información.

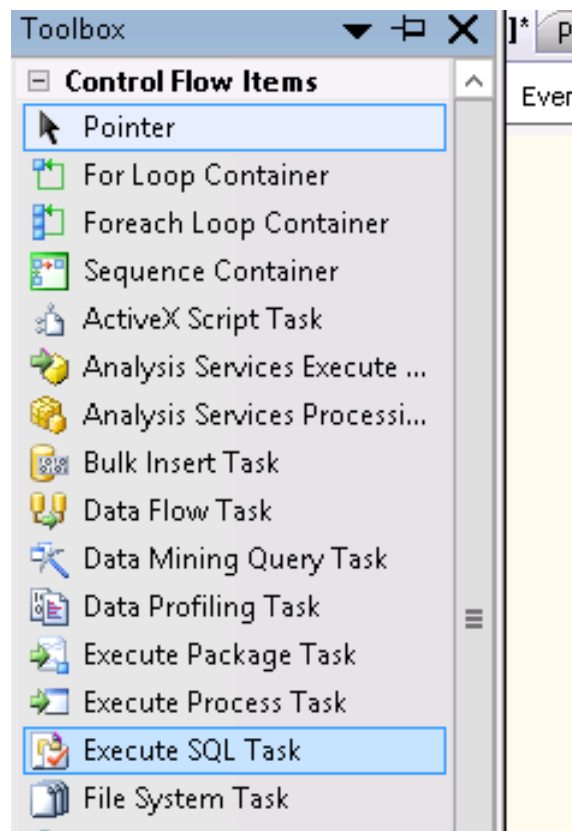


Figura 10: Herramientas del paquete SSIS.

En la pestaña principal *Control Flow*, se tomó el elemento *Execute SQL Task*, que permite hacer consultas directamente a la tabla a través de instrucciones SQL. Al inicio del flujo se elimina la información de datasets anteriores que se encuentra en las tablas (ver Figura 11), debido a que lo que se busca en este proceso es obtener información de un periodo específico, por lo que datos anteriores solo ocasionarían afectarían en los resultados.

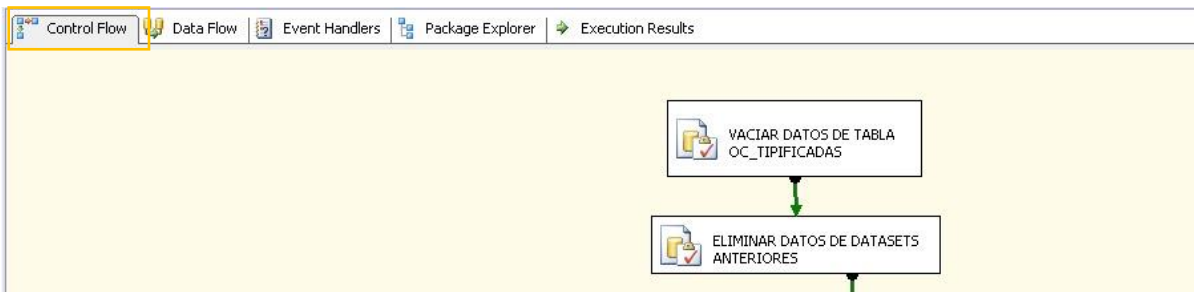


Figura 11: Inicio flujo ETL

Se debe tener en cuenta que dentro de cada elemento se realizó una conexión con la base de datos (Figura 12).

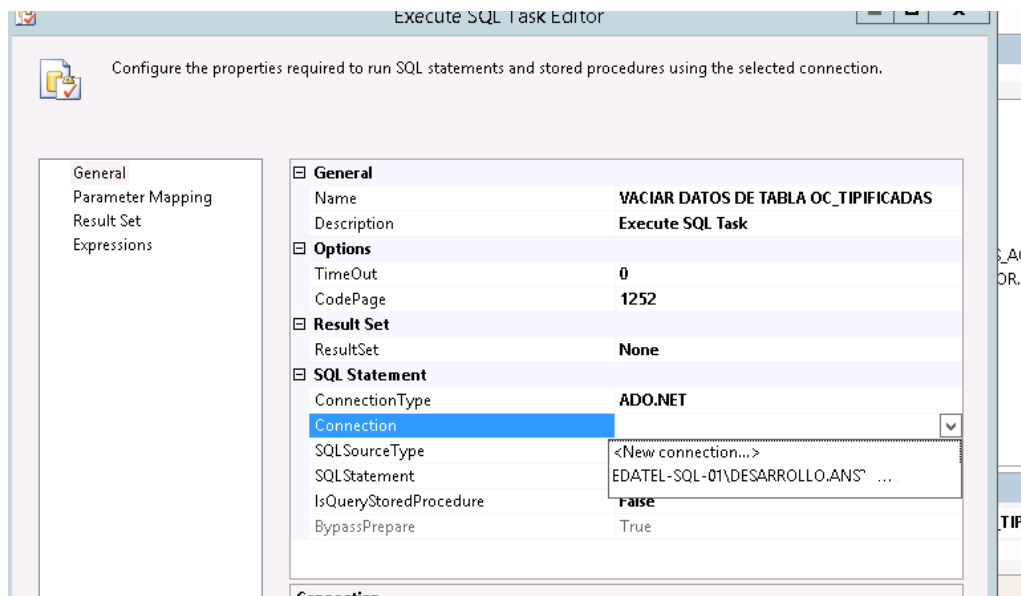


Figura 12. Conexión de cada elemento a la base de datos.

Luego, se formuló la instrucción SQL para ejecutar en la tabla destinada, como se puede apreciar en la Figura 13.

ResultSet	None
SQL Statement	
ConnectionType	ADO.NET
Connection	EDATEL-SQL-01\DESARROLLO.ANSI
SQLSourceType	Direct input
SQLStatement	TRUNCATE TABLE OC_TIPIFICADAS
IsQueryStoredProcedure	False
BypassPrepare	True

Figura 13. Formulación de la instrucción SQL en un elemento.

Los elementos *Data Flow* son utilizados para cargar datos específicos hacia cualquier tabla creada en SSIS desde cualquier archivo como CSV u otra tabla. En el siguiente paso (Figura 14), utiliza ese elemento para cargar la información del archivo `areas_OC_Correctivos_Cerrados_BI_y_Bigdata.csv` a la tabla SDM que se encuentra en SSIS.

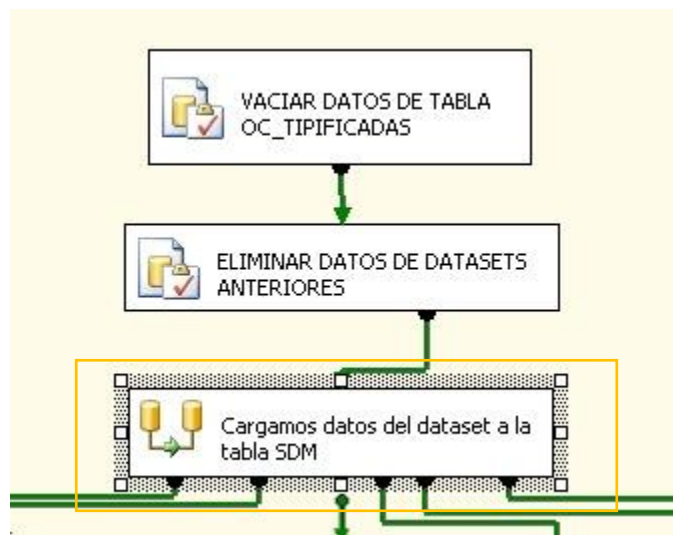


Figura 14. Carga del archivo .CSV a SSIS.

Se utilizaron tres elementos del flujo de datos (ver en la Figura 15):

- **Flat File Source:** Este componente leyó el archivo externo, en este caso, "areas_OC_Correctivos_Cerrados_BI_y_Bigdata.csv".
- **Data Conversion:** Se utilizó para convertir datos que no eran compatibles con las columnas de la tabla SDM. Por ejemplo, los números, que inicialmente eran de tipo cadena y necesitaban convertirse a tipo numérico.
- **ADO NET Destination:** Este componente se encargó de insertar la información en la tabla SDM de SQL Server.

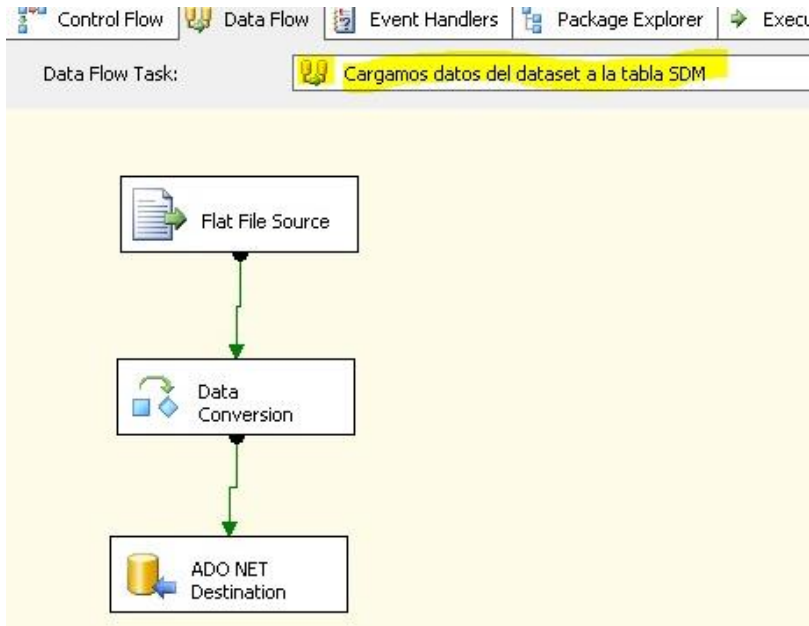


Figura 15. Carga de los datos a la tabla en SSIS.

Para elemento *Flat File Source*, se realizó la conexión al archivo requerido, teniendo en cuenta que su configuración es delimitada por comas y codificación UTF – 8. Esto debe ser indicado en el elemento, como se muestra en la Figura 16.

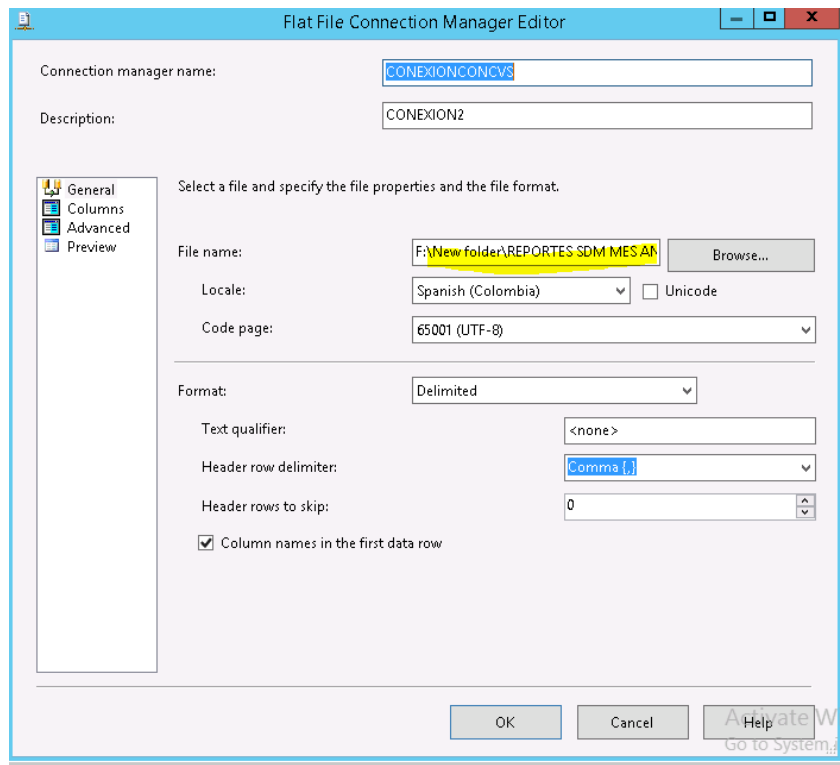
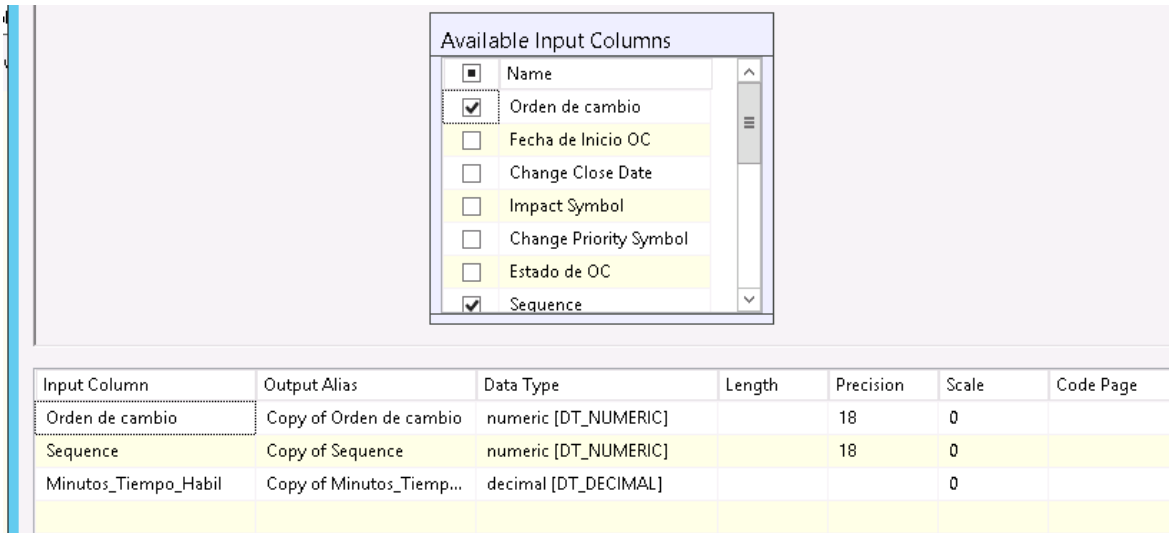


Figura 16. Conexión y configuración de un elemento al archivo .CSV.

A continuación, en el proceso de Data Conversion, se extrajo la información del archivo .CSV, donde los datos estaban en formato de cadena (string). Sin embargo, las columnas de la tabla SDM estaban definidas como numéricas y decimales. Por lo tanto, fue necesario convertir los datos antes de cargarlos en dicha tabla como se muestra en la Figura 17.



Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
Orden de cambio	Copy of Orden de cambio	numeric [DT_NUMERIC]		18	0	
Sequence	Copy of Sequence	numeric [DT_NUMERIC]		18	0	
Minutos_Tiempo_Habil	Copy of Minutos_Tiemp...	decimal [DT_DECIMAL]			0	

Figura 17. *Conversión de las variables.*

En la fase final, se realizó el mapeo de cada columna requerida del dataset CSV a la tabla SDM en SSIS mediante el componente *ADO NET Destination* (ver Figura 18). En este proceso garantiza una transferencia precisa de datos desde el archivo de origen hacia la tabla SSIS, asegurando la integridad de la información y permitiendo una integración fluida entre sistemas de almacenamiento de datos.

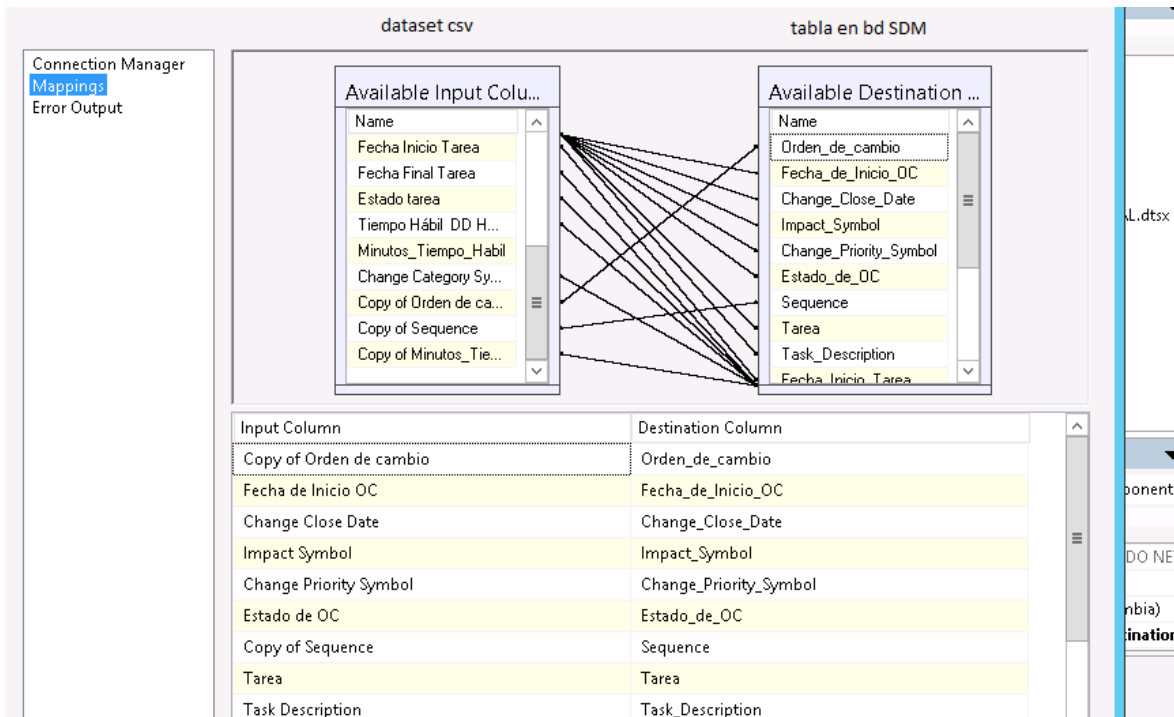


Figura 18. Proceso de mapeo.

Al regresar a la pestaña principal - Control Flow, se incluyeron dos elementos de flujo de datos por categoría. Uno destinado a enviar información a las tablas designadas para las tareas de TIGO, y el otro para las tareas del proveedor (Figura 19). Es importante destacar que se debe seguir un proceso similar al mostrado en Figura 11, Figura 12 y Figura 13, para vaciar las tablas antes de realizar esta carga de datos. En la Tabla 3, se muestran algunas de las consultas en SQL realizadas para este proceso.

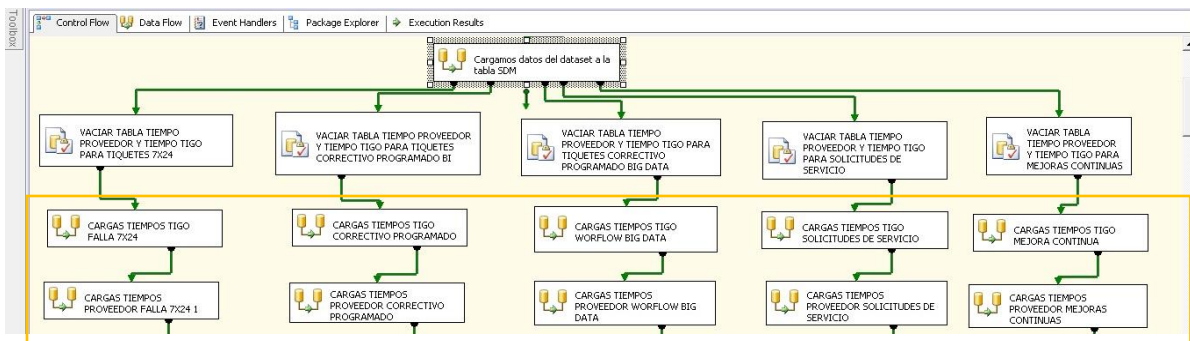


Figura 19. Extracción de datos hacia la tabla de cada tarea y responsable.

Tabla 3. Query para la extracción de datos hacia la tabla de cada tarea y responsable.

Acción	Consulta
Trae la información de impedimentos OC 7x24	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Apl.BI.TT_7x24' and TAREA = 'Impedimento'
Agregar tiempo TIGO correctivo programado	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Apl.Soporte BI.Correctivo Programado' and TAREA IN ('Aval Lider Funcional SOX', 'Aval Lider Tecnico SOX', 'Impedimento', 'Validación CM3 SOX', 'Validación CM2 SOX', 'Validar Artefactos CM2')
Agregar tiempo TIGO solicitudes de servicio	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Catalogo.Configuraciones sistemas corp_Otras_Aplic' and TAREA IN ('Aval Lider Funcional SOX', 'Aval Lider Tecnico SOX', 'Impedimento', 'Validación CM3 SOX', 'Validación CM2 SOX', 'Validar Artefactos CM2')
Agregar tiempo TIGO mejoras continuas	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Catalogo.Plan_Mejora_Continua_Aplicaciones_SOX' and TAREA IN ('Aval Lider Funcional SOX', 'Aval Lider Tecnico SOX', 'Impedimento', 'Validación CM3 SOX', 'Validación CM2 SOX', 'Validar Artefactos CM2') and Sequence <> 82
Agregar tiempo proveedor 7x24	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Apl.BI.TT_7x24' and TAREA <> 'Impedimento'
Agregar tiempo proveedor correctivo programado	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Apl.Soporte BI.Correctivo Programado' and TAREA NOT IN ('Aval Lider Funcional SOX', 'Aval Lider Tecnico SOX', 'Impedimento', 'Validación CM3 SOX',

	'Validación CM2 SOX', 'Validar Artefactos CM2')
Agregar tiempo proveedor solicitudes de servicio	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Catalogo.Configuraciones sistemas corp_Otras_Aplic' and TAREA NOT IN ('Aval Lider Funcional SOX', 'Aval Lider Tecnico SOX', 'Impedimento', 'Validación CM3 SOX', 'Validación CM2 SOX', 'Validar Artefactos CM2')
Agregar tiempo proveedor mejoras continuas	select * from SDM where CHANGE_CATEGORY_SYMBOL = 'Catalogo.Plan_Mejora_Continua_Aplicaciones_SOX' and TAREA NOT IN ('Aval Lider Funcional SOX', 'Aval Lider Tecnico SOX', 'Impedimento', 'Validación CM3 SOX', 'Validación CM2 SOX', 'Validar Artefactos CM2')

Nuevamente, se añadió el elemento *Data Flow* para cargar las OC únicas trabajadas por el proveedor durante el tiempo definido en la tabla OC_tipificadas para cada categoría (Figura 20). Esta acción se llevó a cabo debido a la repetición múltiple del identificador de una OC en la tabla SDM. Esta repetición ocurre porque un mismo ticket puede contener varias tareas, y el enfoque fue trabajar con el tiempo total del ticket en lugar de considerar el tiempo por tarea individualmente. El código utilizado en cada elemento se muestra en la Tabla 4.

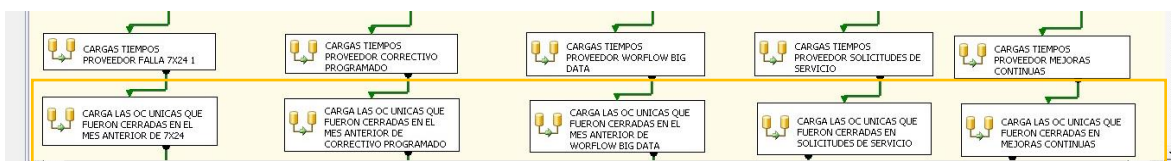


Figura 20. Identificar las OC únicas.

Tabla 4. Query para identificar las OC únicas.

Acción	Consulta
Traer tiquetes únicos 7x24 a la tabla resultante OC tipificada	<pre>SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Falla' as Tipo_de_OC, 'Apl.BI.TT_7x24' as Categoria from dbo.TIEMPO_COMPLETO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion'</pre>
Traer los tiquetes únicos de la fallas, problemas e incidentes de la categoría correctivo programado	<pre>SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Falla' as Tipo_de_OC, 'Apl.Soporte BI.Correctivo Programado' as Categoria from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '1 Alto' AND Change_Priority_Symbol = '4 Alta' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Incidente' as Tipo_de_OC, 'Apl.Soporte BI.Correctivo Programado' as Categoria from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '5 Bajo' AND</pre>

	<p>Change_Priority_Symbol = '2 Baja'</p> <p>UNION ALL</p> <p>SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Problema' as Tipo_de_OC, 'Apl.Soporte BI.Correctivo Programado' as Categoria from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()))-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()))-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '3 Medio' AND Change_Priority_Symbol = '3 Media'</p>
<p>Traer los tiquetes únicos de la fallas, problemas e incidentes de la categoria workflow Big data</p>	<p>SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Falla' as Tipo_de_OC, 'Workflow_Big_Data' as Categoria from dbo.WORFLOW_BIG_DATA WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()))-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()))-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '1 Alto' AND Change_Priority_Symbol = '4 Alta'</p> <p>UNION ALL</p> <p>SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Incidente' as Tipo_de_OC, 'Workflow_Big_Data' as Categoria from dbo.WORFLOW_BIG_DATA WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()))-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()))-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND</p>

	<p>Impact_Symbol = '5 Bajo' AND Change_Priority_Symbol = '2 Baja' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Problema' as Tipo_de_OC, 'Workflow_Big_Data' as Categoria from dbo.WORFLOW_BIG_DATA WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '3 Medio' AND Change_Priority_Symbol = '3 Media'</p>
<p>Traer los tiquetes únicos de las solicitudes de servicio</p>	<p>SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Solicitud de servicio' as Tipo_de_OC, 'Catalogo.Configuraciones sistemas corp_Otras_Aplicaciones' as Categoria from dbo.SOLICITUDES_SERVICIO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-0, 0)) and tarea = 'Gestionar Cierre de la Solicitud'</p>
<p>Traer los tiquetes únicos de las mejoras continuas</p>	<p>SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Mejora Continua' as Tipo_de_OC, 'Catalogo.Plan_Mejora_Continua_Aplicaciones_SOX' as Categoria from dbo.MEJORAS_CONTINUAS WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-0, 0)) and tarea = 'Cierre- Validacion de la Estabilizacion'</p>

En el *Data Flow*, se hizo una carga parcial desde la tabla SDM con el elemento *ADO NET SOURCE* a la tabla de la categoría correspondiente con el elemento *ADO NET DESTINATION*. En este caso, no se necesita un *Data Conversion* porque las columnas están definidas de la misma forma en todas las tablas. La Figura 21 muestra el *Data Flow* que inserta los datos de las tareas realizadas por la empresa de telecomunicaciones en la categoría 7x24. Se realizó un proceso similar por cada categoría y responsable.

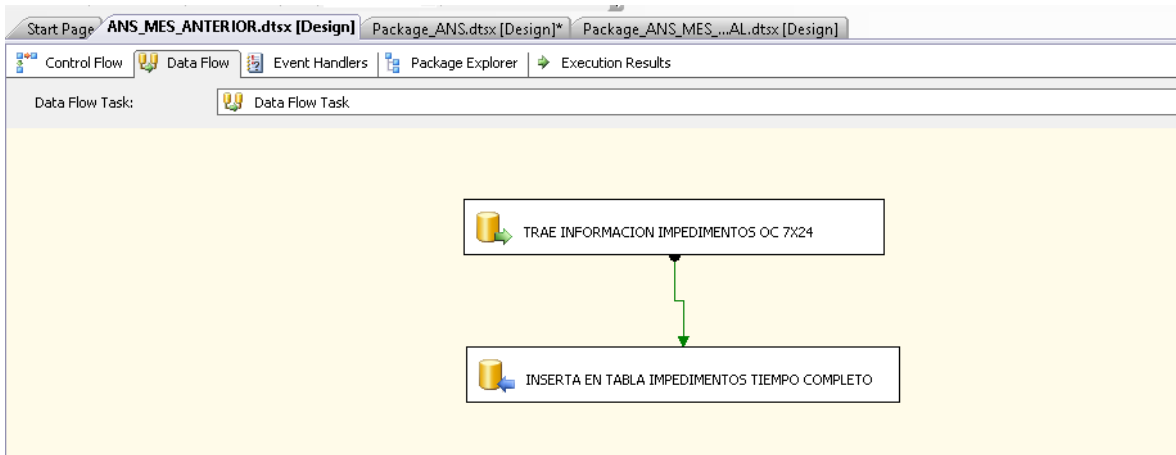


Figura 21. Extracción de la información según el responsable y categoría.

Para concluir el proceso de extracción de datos del archivo original, se excluyó la carga de información de los tickets mal tipificados, es decir, aquellos que presentan datos incorrectos como su nivel de prioridad o de impacto (Figura 22). El código utilizado se muestra en la Tabla 5. Cabe aclarar, que la mala clasificación ocurre debido a que individuos nuevos en el proceso de soporte pueden equivocarse en el tipificar las tareas. Para abordar esta exclusión, se siguieron los mismos pasos previos de vaciado y carga de datos.

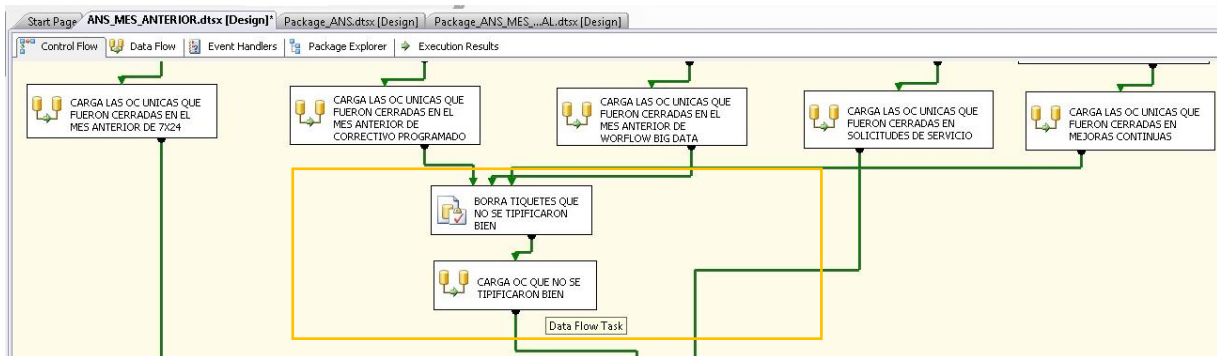


Figura 22. Omisión de la carga de OC mal tipificadas.

Tabla 5. Query omisión de la carga de algunas OC mal tipificadas.

Acción	Consulta
Borrar los datos de la tabla de mal tipificadas	TRUNCATE TABLE OC_MAL_TIIFICADAS
Insertar las OC mal tipificadas - Correctivo programado	<pre> SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Apl.Soporte BI.Correctivo Programado' as Categoria, '1 Alto' AS Impacto, '2 Baja' as Prioridad from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, -1, GETDATE())-1, -1)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '1 Alto' AND Change_Priority_Symbol = '2 Baja' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Apl.Soporte BI.Correctivo Programado' as Categoria, '1 Alto' AS Impacto, '3 Media' as Prioridad from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, -1, GETDATE())-1, -1)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '1 Alto' AND Change_Priority_Symbol = '3 Media' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Apl.Soporte BI.Correctivo Programado' as Categoria, '3 Medio' AS Impacto, '4 Alta' as Prioridad from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, -1, GETDATE())-1, -1)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '3 Medio' AND Change_Priority_Symbol = '4 Alta' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Apl.Soporte BI.Correctivo Programado' as Categoria, '3 Medio' AS Impacto, '2 Baja' as Prioridad from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, </pre>

	<pre> GETDATE()-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '3 Medio' AND Change_Priority_Symbol = '2 Baja' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Apl.Soporte BI.Correctivo Programado' as Categoria, '5 Bajo' AS Impacto, '4 Alta' as Prioridad from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '5 Bajo' AND Change_Priority_Symbol = '4 Alta' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Apl.Soporte BI.Correctivo Programado' as Categoria, '5 Bajo' AS Impacto, '3 Media' as Prioridad from dbo.CORRECTIVO_PROGRAMADO WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '5 Bajo' AND Change_Priority_Symbol = '3 Media' </pre>
<p>Insertar las OC mal tipificadas - workflow Big data</p>	<pre> UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Workflow_Big_Data' as Categoria, '1 Alto' AS Impacto, '2 Baja' as Prioridad from dbo.WORFLOW_BIG_DATA WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE()-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and tarea = 'Cierre- Validacion de la Estabilizacion' AND Impact_Symbol = '1 Alto' AND Change_Priority_Symbol = '2 Baja' UNION ALL SELECT DISTINCT(ORDEN_DE_CAMBIO) AS ORDEN_DE_CAMBIO, 'Workflow_Big_Data' as Categoria, '1 Alto' AS Impacto, '3 Media' as Prioridad from dbo.WORFLOW_BIG_DATA WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, </pre>

```

GETDATE()-1, 0)) AND (select DATEADD(MONTH,
DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and
tarea = 'Cierre- Validacion de la Estabilizacion' AND
Impact_Symbol = '1 Alto' AND
Change_Priority_Symbol = '3 Media'
UNION ALL
SELECT DISTINCT(ORDEN_DE_CAMBIO) AS
ORDEN_DE_CAMBIO, 'Workflow_Big_Data' as
Categoria, '3 Medio' AS Impacto, '4 Alta' as Prioridad
from dbo.WORFLOW_BIG_DATA WHERE
CHANGE_CLOSE_DATE BETWEEN (select
DATEADD(MONTH, DATEDIFF(MONTH, 0,
GETDATE()-1, 0)) AND (select DATEADD(MONTH,
DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and
tarea = 'Cierre- Validacion de la Estabilizacion' AND
Impact_Symbol = '3 Medio' AND
Change_Priority_Symbol = '4 Alta'
UNION ALL
SELECT DISTINCT(ORDEN_DE_CAMBIO) AS
ORDEN_DE_CAMBIO, 'Workflow_Big_Data' as
Categoria, '3 Medio' AS Impacto, '2 Baja' as Prioridad
from dbo.WORFLOW_BIG_DATA WHERE
CHANGE_CLOSE_DATE BETWEEN (select
DATEADD(MONTH, DATEDIFF(MONTH, 0,
GETDATE()-1, 0)) AND (select DATEADD(MONTH,
DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and
tarea = 'Cierre- Validacion de la Estabilizacion' AND
Impact_Symbol = '3 Medio' AND
Change_Priority_Symbol = '2 Baja'
UNION ALL
SELECT DISTINCT(ORDEN_DE_CAMBIO) AS
ORDEN_DE_CAMBIO, 'Workflow_Big_Data' as
Categoria, '5 Bajo' AS Impacto, '4 Alta' as Prioridad
from dbo.WORFLOW_BIG_DATA WHERE
CHANGE_CLOSE_DATE BETWEEN (select
DATEADD(MONTH, DATEDIFF(MONTH, 0,
GETDATE()-1, 0)) AND (select DATEADD(MONTH,
DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and
tarea = 'Cierre- Validacion de la Estabilizacion' AND
Impact_Symbol = '5 Bajo' AND
Change_Priority_Symbol = '4 Alta'
UNION ALL
SELECT DISTINCT(ORDEN_DE_CAMBIO) AS
ORDEN_DE_CAMBIO, 'Workflow_Big_Data' as
Categoria, '5 Bajo' AS Impacto, '3 Media' as Prioridad
from dbo.WORFLOW_BIG_DATA WHERE
CHANGE_CLOSE_DATE BETWEEN (select
DATEADD(MONTH, DATEDIFF(MONTH, 0,
GETDATE()-1, 0)) AND (select DATEADD(MONTH,
DATEDIFF(MONTH, -1, GETDATE()-1, -1)) and
tarea = 'Cierre- Validacion de la Estabilizacion' AND

```

	Impact_Symbol = '5 Bajo' AND Change_Priority_Symbol = '3 Media'
--	--

9.2 Objetivo 2: Transformación de datos

Después de cargar los números de los tiquetes en la tabla OC_Tipificadas, se comenzó a realizar las transformaciones de los datos de las tablas anteriores. Se identificaron las OC que no estaban completadas y donde en la tarea se encontraba la palabra "cancelada", para luego agregar a estos tiquetes el estado de "cancelado". Posteriormente, se buscaron los tiquetes con las fechas de cierre y las tareas completadas, para luego agregarles a estos tiquetes, en la tabla OC_Tipificadas, el estado de "completadas" (ver Figura 23). Esto se hizo por medio de las instrucciones de SQL mostradas en la Tabla 6.

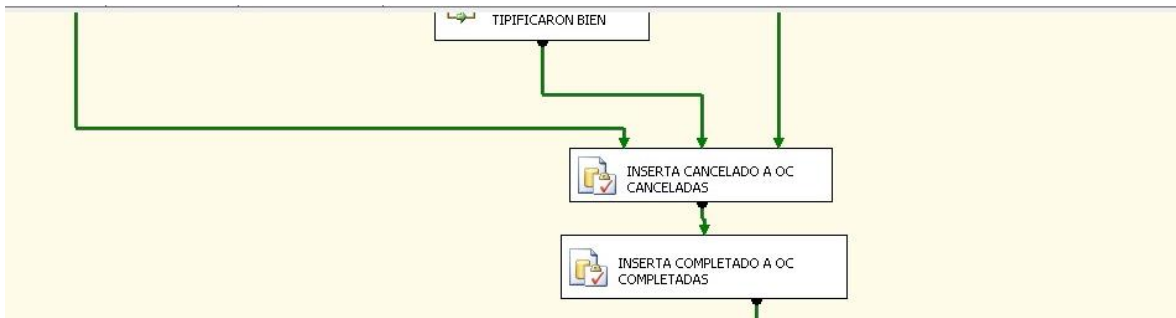


Figura 23. Clasificación de tareas completadas y canceladas.

Tabla 6. Query para la clasificación de tareas completadas y canceladas.

Acción	Código
Hallar las OC canceladas y marcarlas en la tabla de OC tipificadas	<pre> DECLARE @orden_de_camb INT; DECLARE consulta_Cursor CURSOR FOR SELECT ORDEN_DE_CAMBIO FROM SDM where Estado_tarea = 'CNCL'; --SELECT DISTINCT(Orden_de_cambio) as Orden_de_cambio FROM SDM WHERE CHANGE_CLOSE_DATE BETWEEN (select DATEADD(MONTH, DATEDIFF(MONTH, 0, GETDATE())-1, 0)) AND (select DATEADD(MONTH, DATEDIFF(MONTH, - 1, GETDATE())-1, -1)) AND Estado_tarea = 'CNCL' OPEN consulta_Cursor; FETCH consulta_Cursor INTO @orden_de_camb; WHILE @@FETCH_STATUS = 0 BEGIN --PRINT @orden_de_camb UPDATE dbo.OC_TIPIFICADAS SET Estado_OC = 'CANCELADO' WHERE Orden_de_cambio = @orden_de_camb; </pre>

	<pre> FETCH consulta_Cursor INTO @orden_de_camb; END; CLOSE consulta_Cursor; DEALLOCATE consulta_Cursor; </pre>
<p>Marcar las OC que se encuentran completadas en la tabla OC tipificadas</p>	<pre> UPDATE dbo.OC_TIIFICADAS SET Estado_OC = 'COMPLETADA' WHERE Estado_OC IS NULL </pre>

Después, se procedió a calcular los tiempos de las tareas realizadas tanto por el proveedor como por TIGO de forma independiente para cada categoría y número de ticket. Se tuvo en cuenta si se trabajaba con tiempo hábil o tiempo 24/7 (para categorías 7x24 y correctivo programado tipo falla). Una vez calculados, estos tiempos se insertaron en la tabla OC-TIIFICADAS, según el ticket correspondiente en la columna "Tiempo proveedor" o "Tiempo impedimento", como se puede apreciar en la Figura 24. Los cálculos realizados para traer las horas se muestran en la Tabla 7.

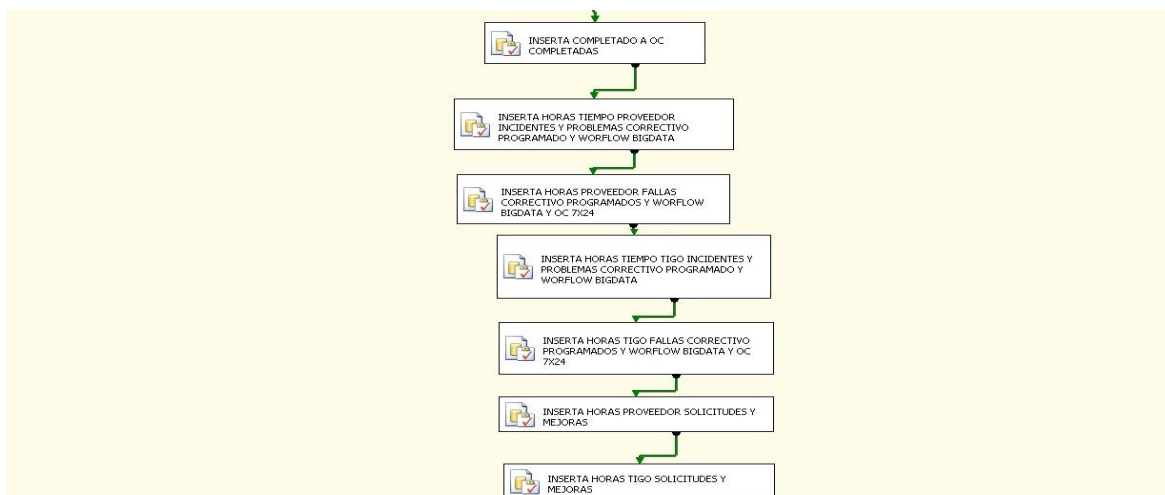


Figura 24. Cálculo de horas trabajadas por responsable y categoría.

Tabla 7. Query para el cálculo de horas trabajadas por responsable y categoría.

Acción	Consulta
<p>Cálculo tiempo total en horas de las fallas de correctivo programdo tanto para BI como Big data</p>	<pre> UPDATE Tabla_A SET Tabla_A.Tiempo_HORAS_trabajadas_SUB = Tabla_B.tiempo_total FROM OC_TIIFICADAS AS Tabla_A INNER JOIN (SELECT (CAST(tiempo_total AS DECIMAL(10))/60) tiempo_total, Orden_de_cambio from (SELECT SUM(c.Tiempo_Trabajado_en_Minutos)tiempo_total, </pre>

	<pre> c.Orden_de_cambio FROM (select b.Orden_de_cambio, DATEDIFF(minute, a.Fecha_Inicio_Tarea, a.Fecha_Final_Tarea) Tiempo_Trabajado_en_Minutos from CORRECTIVO_PROGRAMADO a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio and b.Tipo_de_OC = 'Falla') c group by c.Orden_de_cambio) d UNION ALL SELECT (CAST(tiempo_total AS DECIMAL(10))/60) tiempo_total, Orden_de_cambio from (SELECT SUM(c.Tiempo_Trabajado_en_Minutos)tiempo_total, c.Orden_de_cambio FROM (select b.Orden_de_cambio, DATEDIFF(minute, a.Fecha_Inicio_Tarea, a.Fecha_Final_Tarea) Tiempo_Trabajado_en_Minutos from WORFLOW_BIG_DATA a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio and b.Tipo_de_OC = 'Falla') c group by c.Orden_de_cambio) d UNION ALL SELECT (CAST(tiempo_total AS DECIMAL(10))/60) tiempo_total, Orden_de_cambio FROM (SELECT SUM(c.Tiempo_Trabajado_en_Minutos) tiempo_total, c.ORDEN_DE_CAMBIO FROM (SELECT a.ORDEN_DE_CAMBIO, DATEDIFF(minute, a.Fecha_Inicio_Tarea, a.Fecha_Final_Tarea) Tiempo_Trabajado_en_Minutos FROM TIEMPO_COMPLETO a, OC_TIPIFICADAS b where a.ORDEN_DE_CAMBIO = b.ORDEN_DE_CAMBIO) c group by c.Orden_de_cambio) d) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio </pre>
<p>Cálculo horas de impedimento en problemas e incidentes en correctivo programado para BI y Big data</p>	<pre> UPDATE Tabla_A SET Tabla_A.Tiempo_HORA_impedimentos = Tabla_B.tiempo_total FROM OC_TIPIFICADAS AS Tabla_A INNER JOIN (SELECT (tiempo_total/60) tiempo_total, Orden_de_cambio FROM (select c.tiempo_total, c.Orden_de_cambio from (select SUM(a.Minutos_Tiempo_Habil)tiempo_total, b.Orden_de_cambio from IMPEDIMENTOS_CORRECTIVO_PROGRAMADO a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio group by b.Orden_de_cambio) c, OC_TIPIFICADAS d where c.Orden_de_cambio = d.Orden_de_cambio and Tipo_de_OC IN ('Incidente', 'Problema') UNION ALL select c.tiempo_total, c.Orden_de_cambio from (select SUM(a.Minutos_Tiempo_Habil)tiempo_total, b.Orden_de_cambio from IMPEDIMENTOS_WORFLOW_BIG_DATA a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio group by b.Orden_de_cambio) c, OC_TIPIFICADAS d where c.Orden_de_cambio = </pre>

	<pre>d.Orden_de_cambio and Tipo_de_OC IN ('Incidente', 'Problema')) tabla_virtual_impedimentos_incidentes_y_problemas) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio</pre>
<p>Cálculo de traer la información de fallas tanto correctivo programdo como para OC 7x24 para BI y Big data</p>	<pre>UPDATE Tabla_A SET Tabla_A.Tiempo_HORA_impedimentos = Tabla_B.tiempo_total FROM OC_TIPIFICADAS AS Tabla_A INNER JOIN (SELECT (CAST(tiempo_total AS DECIMAL(10))/60) tiempo_total, Orden_de_cambio from (SELECT SUM(c.Tiempo_Trabajado_en_Minutos)tiempo_total, c.Orden_de_cambio FROM (select b.Orden_de_cambio, DATEDIFF(minute, a.Fecha_Inicio_Tarea, a.Fecha_Final_Tarea) Tiempo_Trabajado_en_Minutos from IMPEDIMENTOS_CORRECTIVO_PROGRAMADO a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio and b.Tipo_de_OC = 'Falla') c group by c.Orden_de_cambio) d UNION ALL SELECT (CAST(tiempo_total AS DECIMAL(10))/60) tiempo_total, Orden_de_cambio from (SELECT SUM(c.Tiempo_Trabajado_en_Minutos)tiempo_total, c.Orden_de_cambio FROM (select b.Orden_de_cambio, DATEDIFF(minute, a.Fecha_Inicio_Tarea, a.Fecha_Final_Tarea) Tiempo_Trabajado_en_Minutos from IMPEDIMENTOS_WORFLOW_BIG_DATA a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio and b.Tipo_de_OC = 'Falla') c group by c.Orden_de_cambio) d UNION ALL SELECT (CAST(tiempo_total AS DECIMAL(10))/60) tiempo_total, Orden_de_cambio FROM (SELECT SUM(c.Tiempo_Trabajado_en_Minutos) tiempo_total, c.ORDEN_DE_CAMBIO FROM (SELECT a.ORDEN_DE_CAMBIO, DATEDIFF(minute, a.Fecha_Inicio_Tarea, a.Fecha_Final_Tarea) Tiempo_Trabajado_en_Minutos FROM IMPEDIMENTOS_TIEMPO_COMPLETO a, OC_TIPIFICADAS b where a.ORDEN_DE_CAMBIO = b.ORDEN_DE_CAMBIO) c group by c.Orden_de_cambio) d) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio</pre>
<p>Cálculo para hallar el tiempo de proveedor para mejores y solicitudes</p>	<pre>UPDATE Tabla_A SET Tabla_A.Tiempo_HORAS_trabajadas_SUB = Tabla_B.tiempo_total FROM</pre>

	<pre> OC_TIPIFICADAS AS Tabla_A INNER JOIN (SELECT (tiempo_total/60) tiempo_total, Orden_de_cambio FROM (select SUM(a.Minutos_Tiempo_Habil)tiempo_total, b.Orden_de_cambio from SOLICITUDES_SERVICIO a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio group by b.Orden_de_cambio UNION ALL select SUM(a.Minutos_Tiempo_Habil)tiempo_total, b.Orden_de_cambio from MEJORAS_CONTINUAS a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio group by b.Orden_de_cambio) tabla_virtual_solicitudes_y_mejoras) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio </pre>
<p>Cálculo para hallar el tiempo de impedimento para mejoras y solicitudes</p>	<pre> UPDATE Tabla_A SET Tabla_A.Tiempo_HORA_impedimentos = Tabla_B.tiempo_total FROM OC_TIPIFICADAS AS Tabla_A INNER JOIN (SELECT (tiempo_total/60) tiempo_total, Orden_de_cambio FROM (select SUM(a.Minutos_Tiempo_Habil)tiempo_total, b.Orden_de_cambio from IMPEDIMENTOS_SOLICITUDES_SERVICIO a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio group by b.Orden_de_cambio UNION ALL select SUM(a.Minutos_Tiempo_Habil)tiempo_total, b.Orden_de_cambio from dbo.IMPEDIMENTOS_MEJORAS_CONTINUAS a, OC_TIPIFICADAS b where a.Orden_de_cambio = b.Orden_de_cambio group by b.Orden_de_cambio) tabla_virtual_impedimentos_incidentes_y_problemas) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio </pre>

Se continuó con la inserción de la fecha de creación y cierre de cada ticket en la tabla de OC_Tipificadas, las cuales fueron tomadas de la tabla SDM (Figura 25). La consulta para la inserción se muestra en la Tabla 8.

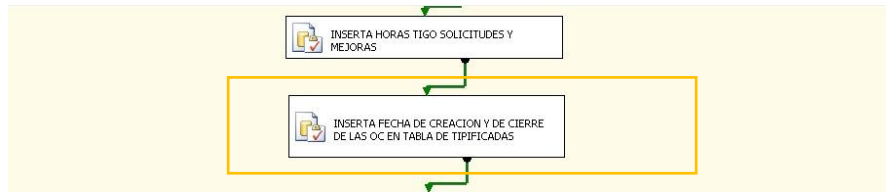


Figura 25. Inserción de la fecha de creación y cierre de cada tickete.

Tabla 8. Query Inserción de la fecha de creación y cierre de cada tickete.

Acción	Consulta
OC tipificadas la fecha de creación y de cierre de cada tickete	<pre> UPDATE Tabla_A SET Tabla_A.Fecha_de_Inicio_OC = Tabla_B.Fecha_de_Inicio_OC, Tabla_A.Fecha_de_Cierre_OC = Tabla_B.Change_Close_Date FROM OC_TIPIFICADAS AS Tabla_A INNER JOIN (Select DISTINCT(A.Orden_de_cambio), A.Fecha_de_Inicio_OC, A.Change_Close_Date from SDM A, OC_TIPIFICADAS B WHERE A.Orden_de_cambio = B.Orden_de_cambio) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio </pre>

Se procedió a calcular las horas desde la fecha de creación y cierre del tickete. Estos resultados se insertaron en la tabla de OC_Tipificadas para el correspondiente identificador del OC (Figura 26). Los cálculos realizados se muestran en la Tabla 9.

Esto se hizo con el fin de realizar la comparación con la suma de tiempo de impedimento y tiempo trabajado por el proveedor.

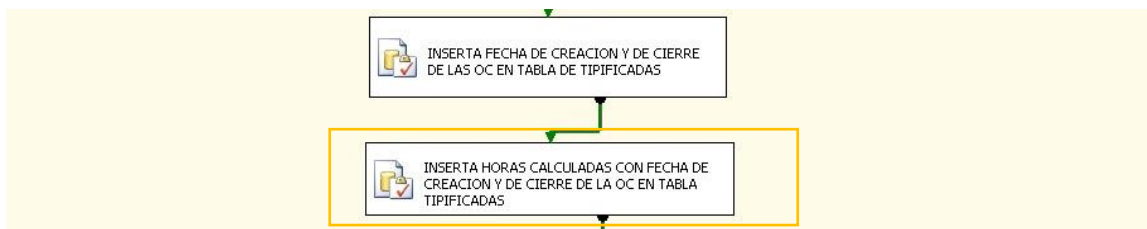


Figura 26. Inserción de las horas calculadas según la fecha de creación y cierre.

Tabla 9. Query para inserción de las horas calculadas según la fecha de creación y cierre

Acción	Consulta
Inserta horas calculadas con fecha de creación y de cierre de la OC en tabla tipificadas	<pre> UPDATE Tabla_A SET Tabla_A.Tiempo_Total_Calulado_con_fecha_Inicio_Cierre_OC = Tabla_B.Tiempo_Calculado_en_horas FROM OC_TIPIFICADAS AS Tabla_A INNER JOIN (SELECT (CAST(Horas_calculadas_con_fechas.Tiempo_Calculado_en_Minutos AS DECIMAL(10))/60) Tiempo_Calculado_en_horas, Horas_calculadas_con_fechas.Orden_de_cambio from (select Orden_de_cambio, DATEDIFF(minute, Fecha_de_Inicio_OC, Fecha_de_Cierre_OC) Tiempo_Calculado_en_Minutos from OC_TIPIFICADAS) Horas_calculadas_con_fechas) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio </pre>

Se sumaron las horas del trabajadas por el proveedor y tiempo de impedimento, y se insertaron los resultados en la tabla OC_TIPIFICADAS, en la columna *Suma tiempos totales*, de acuerdo con el identificador del tickete (Figura 27). El cálculo realizado en SQL se muestra en la Tabla 10.

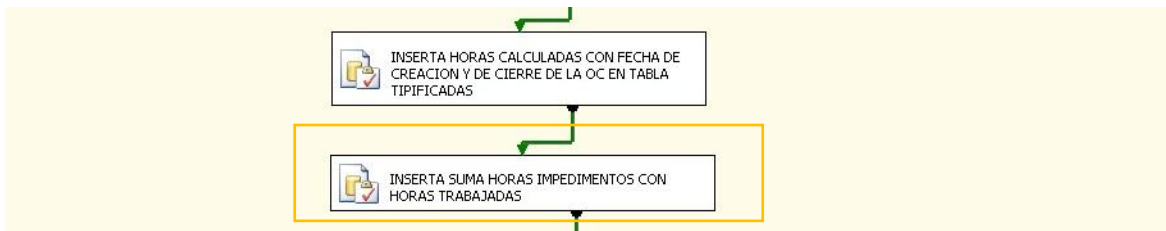


Figura 27. Horas trabajadas totales.

Tabla 10. Query horas trabajadas totales

Acción	Consulta
INSERTA SUMA HORAS IMPEDIMENTO S CON HORAS TRABAJADAS	<pre> UPDATE Tabla_A SET Tabla_A.Suma_Horas_Impedimentos_con_Horas_Trabajadas = Tabla_B.Suma_Horas_Impedimentos_con_Horas_Trabajadas_optenid as FROM OC_TIPIFICADAS AS Tabla_A INNER JOIN (SELECT Orden_de_cambio, </pre>

	<pre> (Tiempo_HORAS_trabajadas_SUB+Tiempo_HORA_impedimentos) Suma_Horas_Impedimentos_con_Horas_Trabajadas_optenidas FROM OC_TIIFICADAS) AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio </pre>
--	---

En este paso, se tomó la prioridad y el impacto para la categoría de mejoras continuas de cada identificador en la tabla SDM y se insertó en la tabla OC tipificadas (Figura 28). En la Tabla 11, se muestra la consulta realizada.

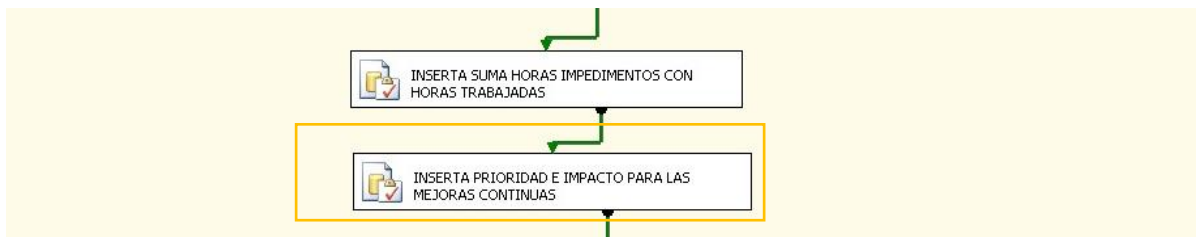


Figura 28. Inserción de la prioridad e impacto para las mejoras continuas.

Tabla 11. Query para inserción de la prioridad e impacto para las mejoras continuas.

Acción	Consulta
Inserta prioridad e impacto para las mejoras continuas	<pre> UPDATE Tabla_A SET Tabla_A.Impact_Symbol = Tabla_B.Impact_Symbol, Tabla_A.Change_Priority_Symbol = Tabla_B.Change_Priority_Symbol FROM OC_TIIFICADAS AS Tabla_A INNER JOIN (SELECT Orden_de_Cambio, Impact_Symbol, Change_Priority_Symbol FROM dbo.MEJORAS_CONTINUAS where Tarea = 'Cierre- Validacion de la Estabilizacion') AS Tabla_B ON Tabla_A.Orden_de_cambio = Tabla_B.Orden_de_cambio </pre>

Para evaluar si el proveedor cumplió con el ANS, se consideran los tiempos de cumplimiento presentados en la Tabla 2.

Si los tiempos del proveedor para cada tiquete son iguales o inferiores a los tiempos de cumplimiento establecidos, se registra "Cumplió" en la columna correspondiente en la tabla OC_tipificadas. De lo contrario, se registra "No cumplió" (ver Figura 29). En código utilizado se puede observar en la Tabla 12.

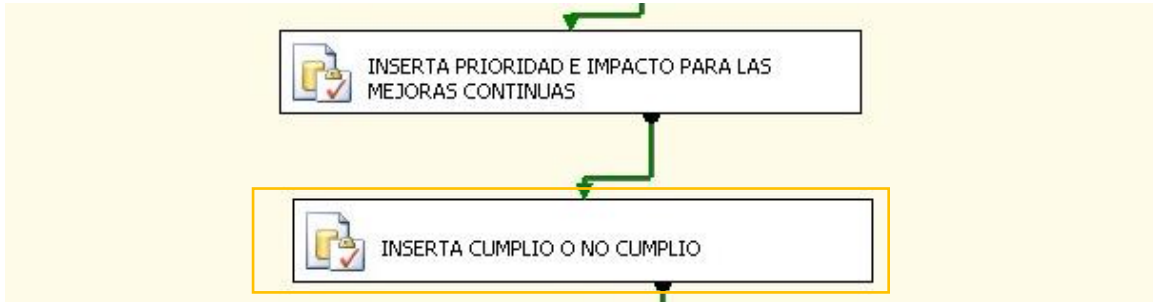


Figura 29. Verificación del cumplimiento del ANS.

Tabla 12. Query para la verificación del cumplimiento del ANS.

Acción	Consulta
Asignación se “cumplió” o “no cumplió”	<pre> update OC_TIIFICADAS SET Cumplimiento = 'CUMPLIO' where Tipo_de_OC = 'Falla' and Tiempo_HORAS_trabajadas_SUB <= 9.5 GO update OC_TIIFICADAS SET Cumplimiento = 'NO CUMPLIO' where Tipo_de_OC = 'Falla' and Tiempo_HORAS_trabajadas_SUB > 9.5 GO update OC_TIIFICADAS SET Cumplimiento = 'CUMPLIO' where Tipo_de_OC = 'Problema' and Tiempo_HORAS_trabajadas_SUB <= 27.5 GO update OC_TIIFICADAS SET Cumplimiento = 'NO CUMPLIO' where Tipo_de_OC = 'Problema' and Tiempo_HORAS_trabajadas_SUB > 27.5 GO update OC_TIIFICADAS SET Cumplimiento = 'CUMPLIO' where Tipo_de_OC = 'Incidente' and Tiempo_HORAS_trabajadas_SUB <= 45.5 GO update OC_TIIFICADAS SET Cumplimiento = 'NO CUMPLIO' where Tipo_de_OC = 'Incidente' and Tiempo_HORAS_trabajadas_SUB > 45.5 GO update OC_TIIFICADAS SET Cumplimiento = 'CUMPLIO' where Tipo_de_OC = 'Solicitud de servicio' and Tiempo_HORAS_trabajadas_SUB <= 18.5 </pre>

	<pre> GO update OC_TIIFICADAS SET Cumplimiento = 'NO CUMPLIO' where Tipo_de_OC = 'Solicitud de servicio' and Tiempo_HORAS_trabajadas_SUB > 18.5 GO update OC_TIIFICADAS SET Cumplimiento = 'CUMPLIO' where Tipo_de_OC = 'Mejora Continua' and Tiempo_HORAS_trabajadas_SUB <= 64 and Impact_Symbol = '5 Bajo' and Change_Priority_Symbol = '2 Baja' GO update OC_TIIFICADAS SET Cumplimiento = 'NO CUMPLIO' where Tipo_de_OC = 'Mejora Continua' and Tiempo_HORAS_trabajadas_SUB > 64 and Impact_Symbol = '5 Bajo' and Change_Priority_Symbol = '2 Baja' GO update OC_TIIFICADAS SET Cumplimiento = 'CUMPLIO' where Tipo_de_OC = 'Mejora Continua' and Tiempo_HORAS_trabajadas_SUB <= 82 and Impact_Symbol = '3 Medio' and Change_Priority_Symbol = '3 Media' GO update OC_TIIFICADAS SET Cumplimiento = 'NO CUMPLIO' where Tipo_de_OC = 'Mejora Continua' and Tiempo_HORAS_trabajadas_SUB > 82 and Impact_Symbol = '3 Medio' and Change_Priority_Symbol = '3 Media' GO update OC_TIIFICADAS SET Cumplimiento = 'CUMPLIO' where Tipo_de_OC = 'Mejora Continua' and Tiempo_HORAS_trabajadas_SUB <= 108 and Impact_Symbol = '1 Alto' and Change_Priority_Symbol = '4 Alta' GO update OC_TIIFICADAS SET Cumplimiento = 'NO CUMPLIO' where Tipo_de_OC = 'Mejora Continua' and Tiempo_HORAS_trabajadas_SUB > 108 and Impact_Symbol = '1 Alto' and Change_Priority_Symbol = '4 Alta' </pre>
--	---

De acuerdo con lo establecido en la Figura 22, no fue necesario dejar en la tabla resultante los tiquetes que fueron mal tipificados, ya que estos no se les podría calcular los cumplimientos y deben de ser ajustados manualmente por la persona que los creó, para luego ser pasados de nuevo por este proceso. Por eso, como se muestra en la Figura 30 y con el código de la Tabla 13, se eliminaron estos registros.

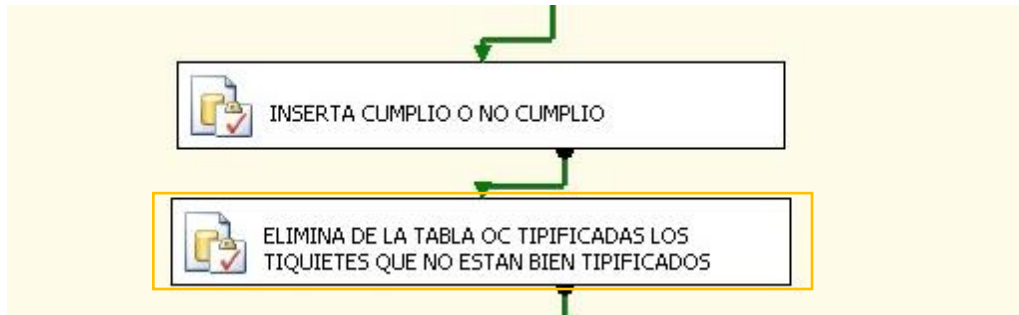


Figura 30. Eliminación de OC mal tipificados.

Tabla 13. Query para la eliminación de OC mal tipificados.

Acción	Consulta
Eliminar de la tabla OC tipificadas los tickets que no están bien tipificados	DELETE FROM dbo.OC_TIPIFICADAS WHERE Orden_de_cambio IN (SELECT Orden_de_cambio FROM OC_MAL_TIPIFICADAS)

9.3 Objetivo 3: Carga de los datos e implementación del modelo

Para la carga de los datos, se generaron tres archivos en Excel a partir de la información de la tabla resultante. Para este proceso, se creó un *Data Flow* (ver Figura 31).

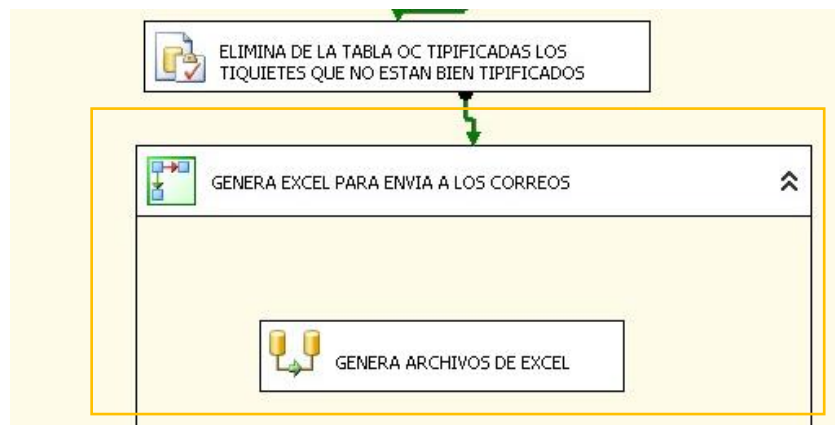


Figura 31. Creación de los archivos en Excel I.

El primer archivo carga los datos de los tickets correctivos programados y fallas, los cuales pertenecen a los tickets de soporte. El segundo archivo carga los datos de las mejoras y solicitudes de servicio, los cuales pertenecen a desarrollos menores. Por último, el tercer archivo, carga la información de los tickets mal tipificados. Lo anterior se realizó

con un *ADO NET SOURCE*. La creación del archivo Excel se realizó con un *FLAT FILE DESTINATION* (ver Figura 32).

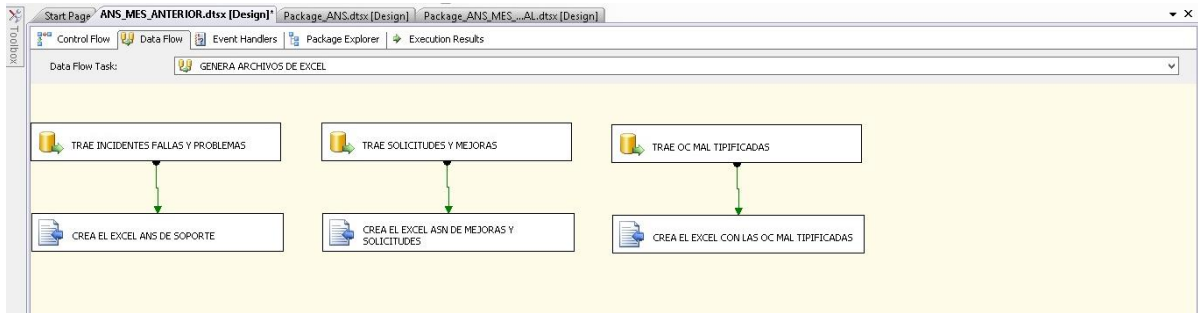


Figura 32. Creación de los archivos en Excel II.

Los archivos generados se guardan directamente en la ruta indicada en el servidor (Figura 33). Para, posteriormente, ser adjuntados y enviados por correo.

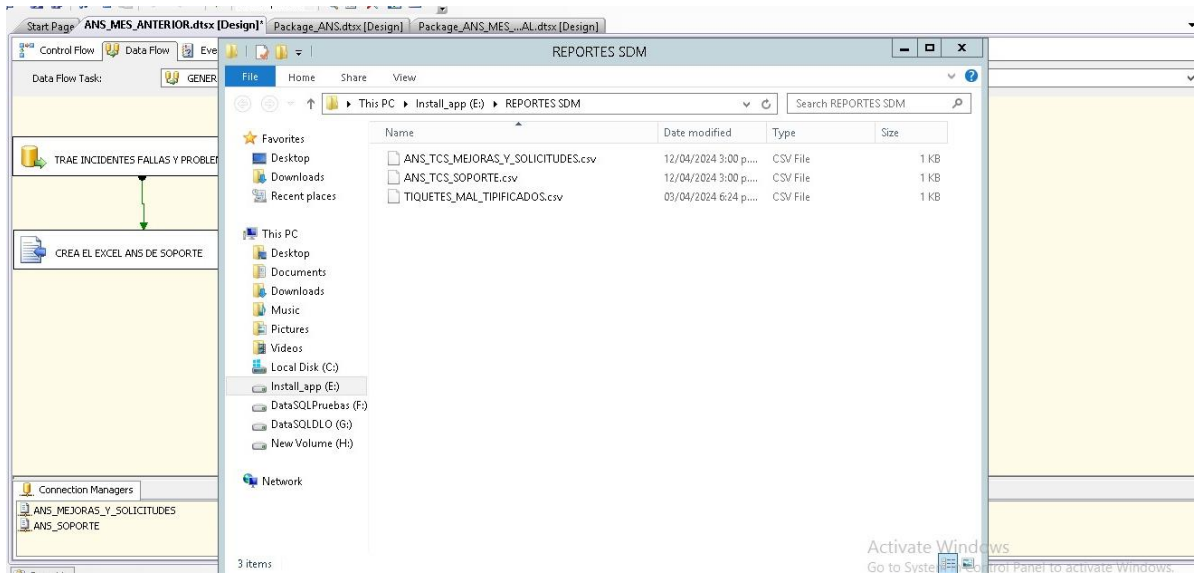


Figura 33. Archivos guardados en la ruta definida.

Para realizar el envío por correo electrónico, se utilizó el elemento *SEND MAIL TASK* (Figura 34)

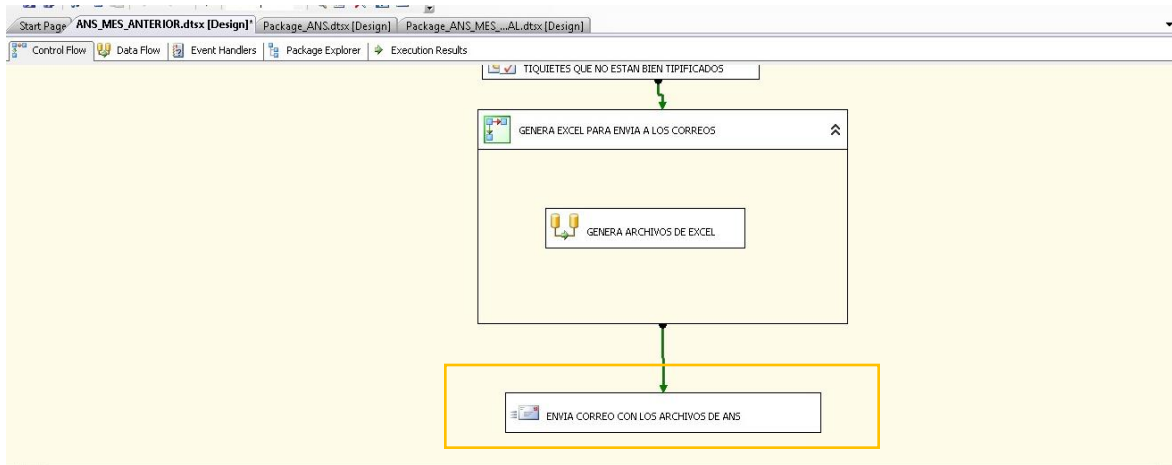


Figura 34. Elemento para el envío del correo electrónico.

En la Figura 35, se muestra la configuración para el elemento *SEND MAIL TASK*, donde se indicó el email del destinatario, asunto y cuerpo del correo, y se adjuntaron los archivos requeridos.

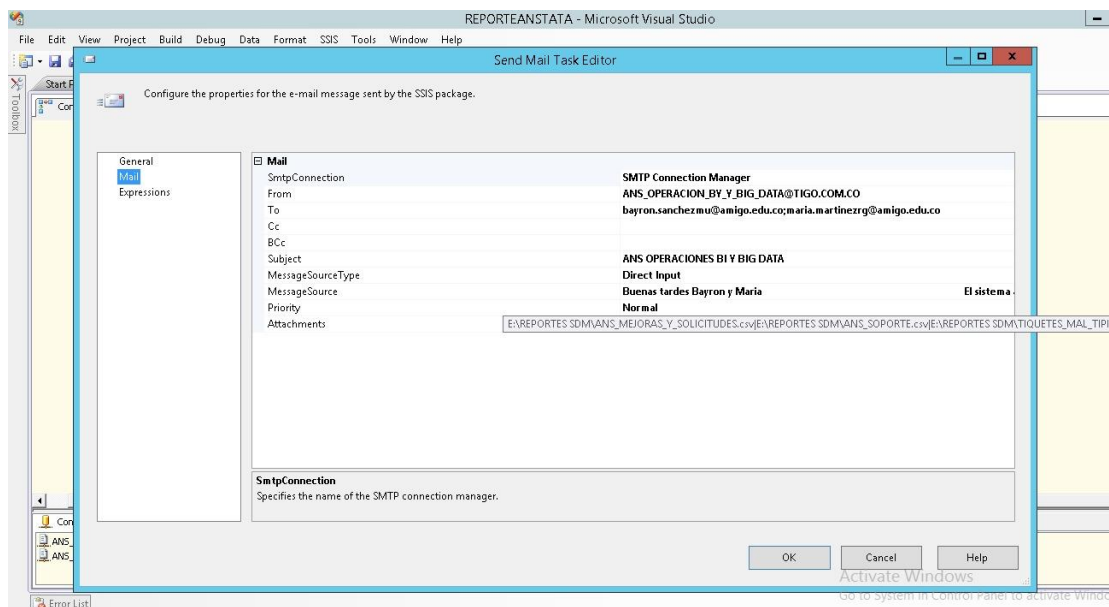


Figura 35. Configuración del elemento para el envío del correo.

En la Figura 36 se muestra un ejemplo de cómo se recibe el correo enviado por SSIS.

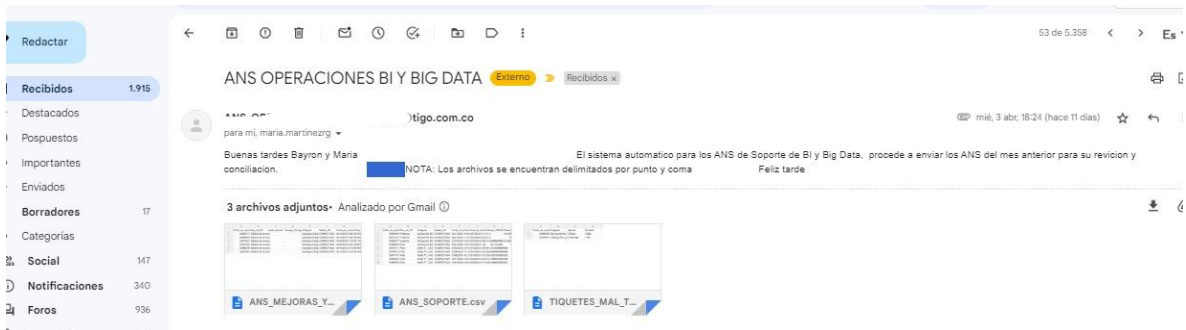


Figura 36. Correo final.

El propósito de implementar este modelo era generar tres reportes finales para el usuario, cada uno conteniendo información correspondiente a un periodo específico. Por este motivo, se realiza una limpieza inicial de los datos y se trabaja con un rango determinado. Aunque el modelo tiene la capacidad de manejar un mayor volumen de datos, se establecen estos límites debido a las particularidades del negocio. Sin embargo, incluso si se generara un gran número de tiquetes y tareas, no se vería afectado el rendimiento del modelo en ningún momento.

9.4 Objetivo general: Automatización en SSIS

La ejecución del ETL no debe ser manual, ya que el usuario tendría que entrar al servidor y ejecutarlo, lo cual no es posible debido a las políticas de seguridad de la empresa. Además, se buscaba facilitar el trabajo de los usuarios. Por esto, se crearon dos Jobs para que ejecuten cada ETL (Figura 37), uno del mes actual y otro del mes anterior.

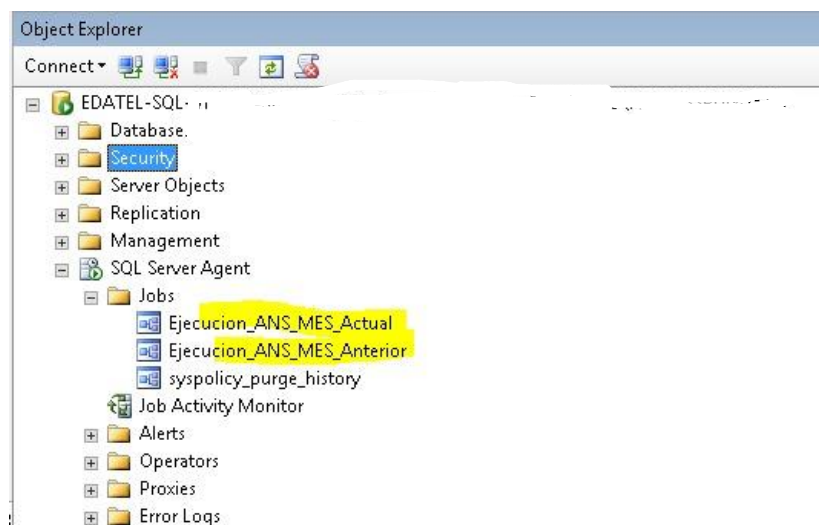


Figura 37. Creación de Jobs.

Para el “mes actual” se ejecuta una vez por semana durante el respectivo mes, generalmente los viernes. Se realizó la configuración de acuerdo con esta necesidad como se puede observar en la Figura 38.

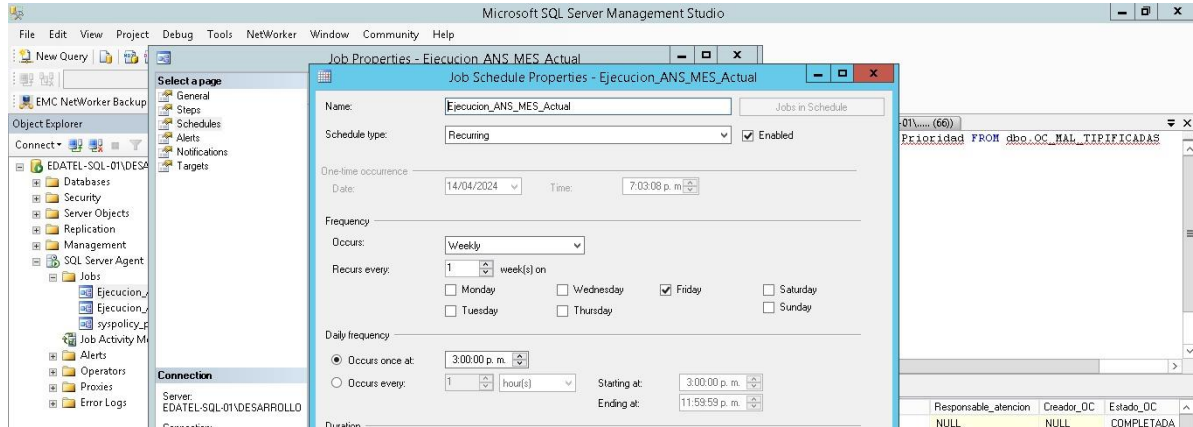


Figura 38: Configuración Job para el mes actual.

En la Figura 39, se visualizan los reprocesos desde el mes de marzo, lo que proporciona una muestra de las pruebas realizadas con este Job.

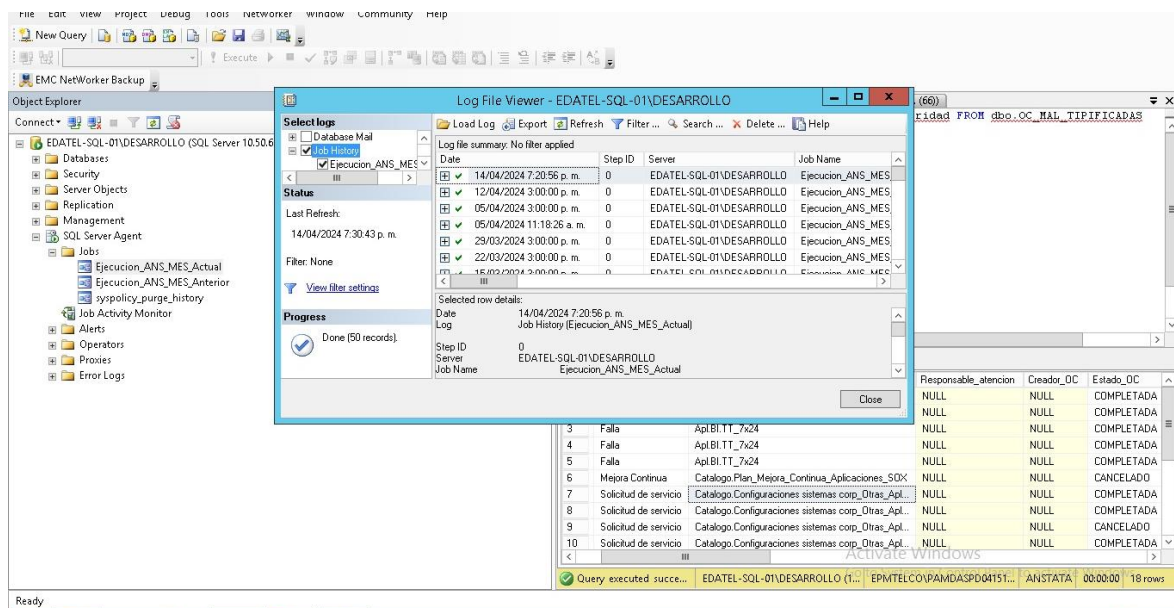


Figura 39: Reprocesamiento para el mes actual.

Para el mes anterior, se ejecuta el Job el segundo día de cada mes debido a que en el primer día se carga el archivo. En la Figura 40 se muestra la configuración de este Job.

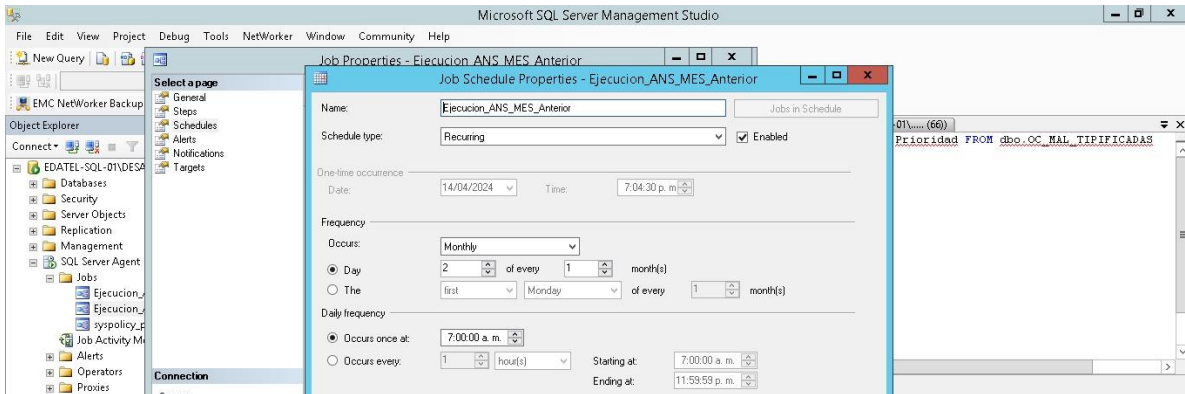


Figura 40. Configuración Job del mes anterior.

En la Figura 41, se visualizan los reprocesos desde el mes de marzo, lo que proporciona una muestra de las pruebas realizadas con este Job.

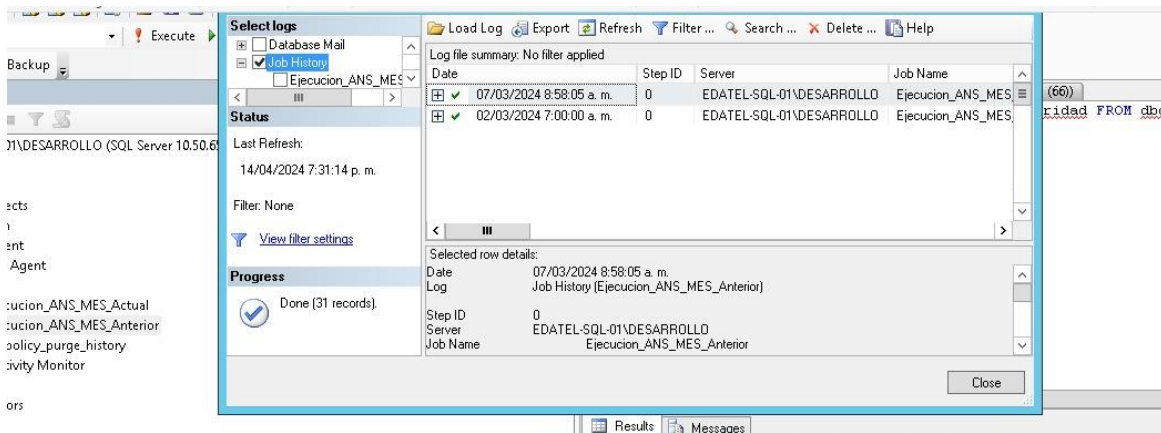


Figura 41. Reprocesamiento para el mes actual.

La automatización del proceso ETL ha logrado reducir drásticamente el tiempo dedicado a la generación y envío de los informes, ahora completado en menos de 5 minutos. Anteriormente, este trabajo implicaba aproximadamente 4.5 horas de labor diaria por parte de tres personas. Este resultado evidencia una optimización sustancial del proceso, liberando a los trabajadores para que puedan enfocarse en tareas más significativas para el negocio.

9.5 Resultado adicional: Inteligencia de negocios

Además, como resultado adicional del proyecto, se logró proporcionar información completa y precisa para facilitar la inteligencia de negocios. Mediante la automatización del proceso ETL, se pudo extraer los datos relacionados del ANS de manera rápida y eficiente para múltiples meses, lo que permitió la creación de dos gráficos para analizar el

rendimiento mensual en función de las horas de impedimento y el proveedor, según la categoría de las tareas.

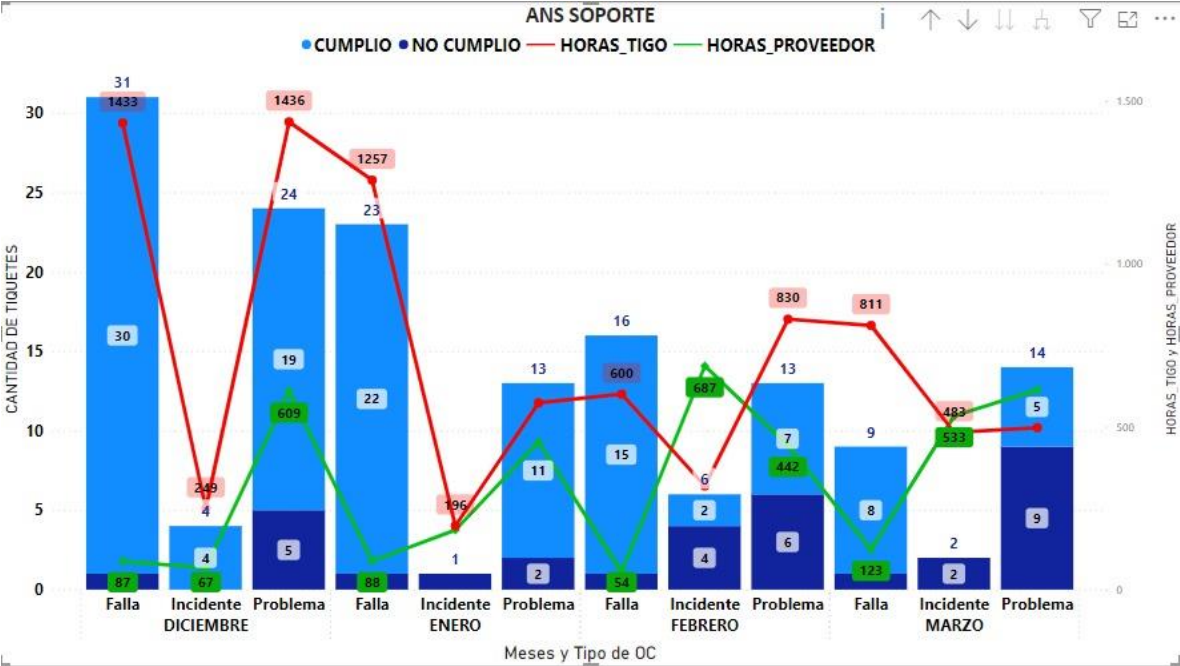


Figura 42. ANS fallas, incidentes y problemas.

El desempeño del proveedor y de la empresa de telecomunicaciones al realizar tareas de soporte (fallas, incidentes y problemas) se muestra en la Figura 42. Las barras representan la cantidad de órdenes que cumplieron con los ANS y las que no. Las líneas muestran las horas dedicadas por cada responsable durante la ejecución de cada orden.

Se observa que, en general, se reportan más fallas durante los meses analizados, mientras que los incidentes presentan una menor incidencia, pero una mayor cantidad de incumplimientos en comparación. Por otro lado, se destaca que TIGO requiere una considerable cantidad de tiempo para completar las órdenes, superando el tiempo empleado por el proveedor. Este aspecto requiere mejora, ya que la externalización de este servicio tenía como objetivo principal reducir la carga laboral interna. Sin embargo, en general, se observa un buen cumplimiento en estas órdenes de servicio.

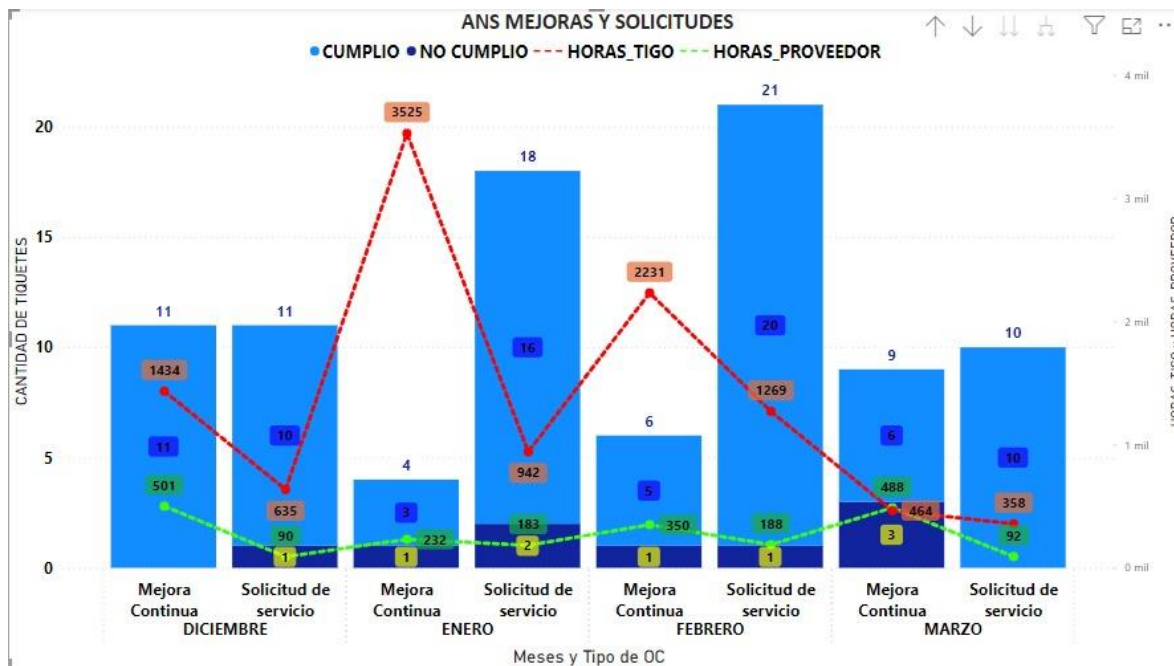


Figura 43. ANS mejoras continuas y solicitudes de servicio.

Aunque las tareas de desarrollos menores, como mejoras continuas y solicitudes de servicio, muestran una demanda menor en comparación con el grupo anterior, como se evidencia en la Figura 43, observamos un comportamiento similar en términos de tiempos de impedimento, especialmente en el caso de mejoras continuas. Durante los meses de enero y febrero, se registró un aumento significativo en las horas dedicadas por TIGO, siendo aproximadamente un 1419% más en enero y un 537% más en febrero en comparación con las horas empleadas por el subcontratista para estas OC. Como se mencionó anteriormente, es fundamental identificar las causas detrás de estas horas adicionales empleadas por la empresa de telecomunicaciones y buscar formas de reducirlas. Es relevante destacar que, en marzo, para una sola categoría, el contratista empleó un 5.2% más de horas que en meses anteriores, lo que subraya la necesidad de una gestión más eficiente de los recursos en esta área.

10. Conclusiones

Se demostró la viabilidad de optimizar el proceso mediante la aplicación de técnicas de minería de datos en Integration Services y la programación con SQL. Este logro trajo beneficios significativos para la empresa, tanto en la reducción del tiempo empleado para la verificación del cumplimiento de los ANS como en el aumento de la calidad de los resultados mediante la automatización del proceso ETL, por lo que se pudo responder adecuadamente la pregunta de investigación general del trabajo.

Los objetivos planteados fueron alcanzados satisfactoriamente. Se extrajo la información del trabajo realizado por el subcontratista del sistema de tiquetes de gestión “servicedesk” de la empresa, se transformaron los datos recopilados de acuerdo con los tipos de ANS, y se cargó la información requerida para el envío por correo electrónico de los informes del ANS al área responsable. Además, se automatizó este proceso, eliminando la necesidad de intervención directa por parte del personal.

Adicionalmente, se logró un alcance extra con la generación de reportes visuales, lo cual facilitó la toma de decisiones en la empresa. Este logro representa un valor añadido tanto para este proyecto como para la empresa, ya que al contar con estos tableros que se alimentan automáticamente de los reportes, las áreas interesadas pueden ver en tiempo real el desempeño del subcontratista y de la empresa. Esto permite identificar oportunidades de mejora de manera oportuna.

11. Recomendaciones

Es esencial realizar comparativas que permitan comprender las ventajas reales de la implementación de estos proyectos, específicamente en lo que respecta a la automatización del proceso de ETL. Para ello, es necesario recopilar datos históricos que reflejen el estado actual del proceso en términos de costos y tiempo. Posteriormente, estas mismas métricas deberían ser registradas después de la implementación del proyecto. Los artículos revisados no proporcionaron este tipo de métricas comparativas.

Por otro lado, las propuestas revisadas en los artículos consultados para los antecedentes y en el presente proyecto, se basaron en el uso de software con licencia paga. Sería interesante explorar la posibilidad de llevar a cabo proyectos similares utilizando software de código abierto. Esto permitiría realizar comparaciones en términos de problemas encontrados, tiempo de desarrollo, facilidad de uso, costos y otras variables relevantes.

Adicionalmente, existe un amplio abanico de herramientas que pueden mejorar este tipo de proyectos, lo que abre un mundo de posibilidades. Por ejemplo, se puede considerar la utilización de técnicas de machine-learning para generar indicadores predictivos, lo que facilita la toma de decisiones en la organización. Esta herramienta también podría resultar útil en la gestión de bases de datos, especialmente cuando se trata de Big Data, ya que el formato de fuente de datos podría cambiar y afectar el proceso automatizado de ETL. La implementación de machine-learning permitiría adaptar estos cambios sin necesidad de intervención manual, lo que optimizaría el tiempo y reduciría la posibilidad de errores.

12. Referencias

- Behnke, M., Valik, J. K., Gubbels, S., Teixeira, D., Kristensen, B., Abbas, M., van Rooden, S. M., Gastmeier, P., van Mourik, M. S. M., Aspevall, O., Astagneau, P., Bonten, M. J. M., Carrara, E., Gomila-Grange, A., de Greeff, S. C., Harrison, W., Humphreys, H., Johansson, A., Koek, M. B. G., ... Tängdén, T. (2021). Information technology aspects of large-scale implementation of automated surveillance of healthcare-associated infections. *Clinical Microbiology and Infection*, 27, S29–S39. <https://doi.org/10.1016/j.cmi.2021.02.027>
- Boulila, W., Al-kmal, M., Farid, M., & Mugahed, H. (2023). A business intelligence based solution to support academic affairs: case of Taibah University. *Wireless Networks*, 29(3), 1051–1058. <https://doi.org/10.1007/s11276-018-1880-3>
- Bunod, R., Augstburger, E., Brasnu, E., Labbe, A., & Baudouin, C. (2022). Artificial Intelligence and Business Value: A literature review. *Journal Francais d’Ophthalmologie*, 45(2), 216–232. <https://doi.org/10.1016/j.jfo.2021.11.002>
- Coito, T., Martins, M. S. E., Viegas, J. L., Firme, B., Figueiredo, J., Vieira, S. M., & Sousa, J. M. C. (2020). A Middleware Platform for Intelligent Automation: An Industrial Prototype Implementation. *Computers in Industry*, 123, 103329. <https://doi.org/https://doi.org/10.1016/j.compind.2020.103329>
- Deist, T. M., Dankers, F. J. W. M., Ojha, P., Scott Marshall, M., Janssen, T., Faivre-Finn, C., Masciocchi, C., Valentini, V., Wang, J., Chen, J., Zhang, Z., Spezi, E., Button, M., Jan Nuyttens, J., Vernhout, R., van Soest, J., Jochems, A., Monshouwer, R., Bussink, J., ... Dekker, A. (2020). Distributed learning on 20 000+ lung cancer patients – The Personal Health Train. *Radiotherapy and Oncology*, 144, 189–200. <https://doi.org/10.1016/j.radonc.2019.11.019>
- Dogea, R., & Stolt, R. (2021). Identifying Challenges Related to Industry 4.0 in Five Manufacturing Companies. *Procedia Manufacturing*, 55, 328–334. <https://doi.org/https://doi.org/10.1016/j.promfg.2021.10.046>
- Función Pública - Gobierno de Colombia. (n.d.). *Glosario - Acuerdo de Nivel de Servicio (ANS)*. Retrieved April 17, 2024, from [https://www.funcionpublica.gov.co/glosario/-/wiki/26415658/Acuerdo de Nivel de Servicio \(ANS\)#:~:text=Los Acuerdos de Niveles de,medirá la oportunidad del mismo.](https://www.funcionpublica.gov.co/glosario/-/wiki/26415658/Acuerdo+de+Nivel+de+Servicio+(ANS)#:~:text=Los+Acuerdos+de+Niveles+de,medirá+la+oportunidad+del+mismo.)
- Gaardboe, R., Nyvang, T., & Sandalgaard, N. (2017). Business Intelligence Success applied to Healthcare Information Systems. *Procedia Computer Science*, 121, 483–490. <https://doi.org/10.1016/j.procs.2017.11.065>

- Hmoud, H., Al-Adwan, A. S., Horani, O., Yaseen, H., & Zoubi, J. Z. Al. (2023). Factors influencing business intelligence adoption by higher education institutions. *Journal of Open Innovation: Technology, Market, and Complexity*, 9(3), 100111. <https://doi.org/10.1016/j.joitmc.2023.100111>
- IBM. (n.d.-a). *¿Qué es la calidad de datos?* Retrieved September 30, 2023, from <https://www.ibm.com/mx-es/topics/what-is-data-quality>
- IBM. (n.d.-b). *¿Qué es la gestión de Tecnologías de la Información?* Retrieved September 30, 2023, from <https://www.ibm.com/mx-es/topics/it-management>
- IBM. (n.d.-c). *¿Qué es la inteligencia artificial (IA)?* Retrieved September 29, 2023, from <https://www.ibm.com/mx-es/topics/artificial-intelligence>
- Keskin, S., & Yazıcı, A. (2023). FSOLAP: A fuzzy logic-based spatial OLAP framework for effective predictive analytics. *Expert Systems with Applications*, 213(Part B). <https://doi.org/10.1016/j.eswa.2022.118961>
- Kubina, M., Koman, G., & Kubinova, I. (2015). Possibility of Improving Efficiency within Business Intelligence Systems in Companies. *Procedia Economics and Finance*, 26(15), 300–305. [https://doi.org/10.1016/s2212-5671\(15\)00856-4](https://doi.org/10.1016/s2212-5671(15)00856-4)
- Llave, M. R. (2017). Business Intelligence and Analytics in Small and Medium-sized Enterprises: A Systematic Literature Review. *Procedia Computer Science*, 121, 194–205. <https://doi.org/10.1016/j.procs.2017.11.027>
- Machado, G. V., Cunha, Í., Pereira, A. C. M., & Oliveira, L. B. (2019). DOD-ETL: distributed on-demand ETL for near real-time business intelligence. *Journal of Internet Services and Applications*, 10(1), 1–15. <https://doi.org/10.1186/s13174-019-0121-z>
- Medina, F., & Galván, M. (2007). Imputación de datos: teoría y práctica. In NU. CEPAL. División de Estadística y Proyecciones Económicas (Ed.), *Estudios estadísticos y prospectivos* (Vol. 4). NACIONES UNIDAS. <https://doi.org/978-92-1-323101-2>
- Microsoft. (2021a). *Access SQL: conceptos básicos, vocabulario y sintaxis*. <https://support.microsoft.com/es-es/office/access-sql-conceptos-básicos-vocabulario-y-sintaxis-444d0303-cde1-424e-9a74-e8dc3e460671>
- Microsoft. (2021b). *Información general sobre el procesamiento analítico en línea (OLAP)*. [https://support.microsoft.com/es-es/office/información-general-sobre-el-procesamiento-analítico-en-línea-olap-15d2cdde-f70b-4277-b009-ed732b75fdd6#:~:text=El procesamiento analítico en línea \(OLAP\) es una tecnología que,grandes y admitir inteligencia empr](https://support.microsoft.com/es-es/office/información-general-sobre-el-procesamiento-analítico-en-línea-olap-15d2cdde-f70b-4277-b009-ed732b75fdd6#:~:text=El procesamiento analítico en línea (OLAP) es una tecnología que,grandes y admitir inteligencia empr)

- Microsoft. (2023). *SQL Server Integration Services*. <https://learn.microsoft.com/es-es/sql/integration-services/sql-server-integration-services?view=sql-server-ver16>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *88*, 105906. <https://doi.org/https://doi.org/10.1016/j.ijisu.2021.105906>
- Power, D. J., Heavin, C., McDermott, J., & Daly, M. (2018). Defining business analytics: an empirical approach. *Journal of Business Analytics*, *1*(1), 40–53. <https://doi.org/10.1080/2573234X.2018.1507605>
- Schmied, S., Mathias, S. G., Großmann, D., & Jumar, U. (2020). Integration of existing cyber-physical manufacturing systems into a common information model. *IFAC-PapersOnLine*, *53*(2), 10905–10910. <https://doi.org/10.1016/j.ifacol.2020.12.2827>
- Schuh, G., Reinhart, G., Prote, J.-P., Sauermann, F., Horsthofer, J., Oppolzer, F., & Knoll, D. (2019). Data Mining Definitions and Applications for the Management of Production Complexity. *Procedia CIRP*, *81*, 874–879. <https://doi.org/https://doi.org/10.1016/j.procir.2019.03.217>
- Seibold, J., Hentsch, M., Kharitonov, A., Eber, R., & Schwarzer, S. (2022). Process automation in the area of manufacturability analysis using machine learning. *Procedia Computer Science*, *204*, 196–204. <https://doi.org/10.1016/j.procs.2022.08.024>
- Souibgui, M., Atigui, F., Zammali, S., Cherfi, S., & Yahia, S. Ben. (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>
- TIGO Colombia. (2023). *Conócenos, somos TIGO Colombia*. <https://www.TIGO.com.co/conocenos/quien-es-TIGO>
- Václav, C., Gabriel, F., Blanka, K., Libor, K., & Michal, T. (2021). Utilization of business intelligence tools in cargo control. *Transportation Research Procedia*, *53*(2019), 212–223. <https://doi.org/10.1016/j.trpro.2021.02.028>
- Velásquez, J. D. (2015). Una guía corta para escribir revisiones sistemáticas de literatura parte 3. *DYNA (Colombia)*, *82*(189), 9–12. <https://doi.org/10.15446/dyna.v82n189.48931>
- Wu, S. S., Shirley, R. B., Anne, S., Georgopoulos, R., Appachi, S., & Hopkins, B. (2022). Utility of the finance-electronic medical record digital dashboard in pediatric

otolaryngology. *American Journal of Otolaryngology - Head and Neck Medicine and Surgery*, 43(5). <https://doi.org/10.1016/j.amjoto.2022.103598>