

PRODUCTO CLOUD B2B: UNIFICACIÓN Y OPTIMIZACIÓN DEL PROCESO DE GESTIÓN DE LA INFORMACIÓN DEL PRODUCTO CLOUD DE TIGO BUSINESS

PRACTICANTE: Stefany Carolina De Sousa Lora

PROGRAMA: Ingeniería de Sistemas

ASESORES: Luis Hernando Silva Florez, Luis Fernando Suarez Giraldo

Semestre de la práctica: 2024-1

Tigo es una empresa de telecomunicaciones radicada en Colombia que ofrece una amplia variedad de servicios tanto a clientes como empresas. Entre los servicios ofrecidos se encuentra la telefonía fija y móvil, internet fijo, televisión por cable o satélite, entre otras soluciones digitales de índole corporativa.



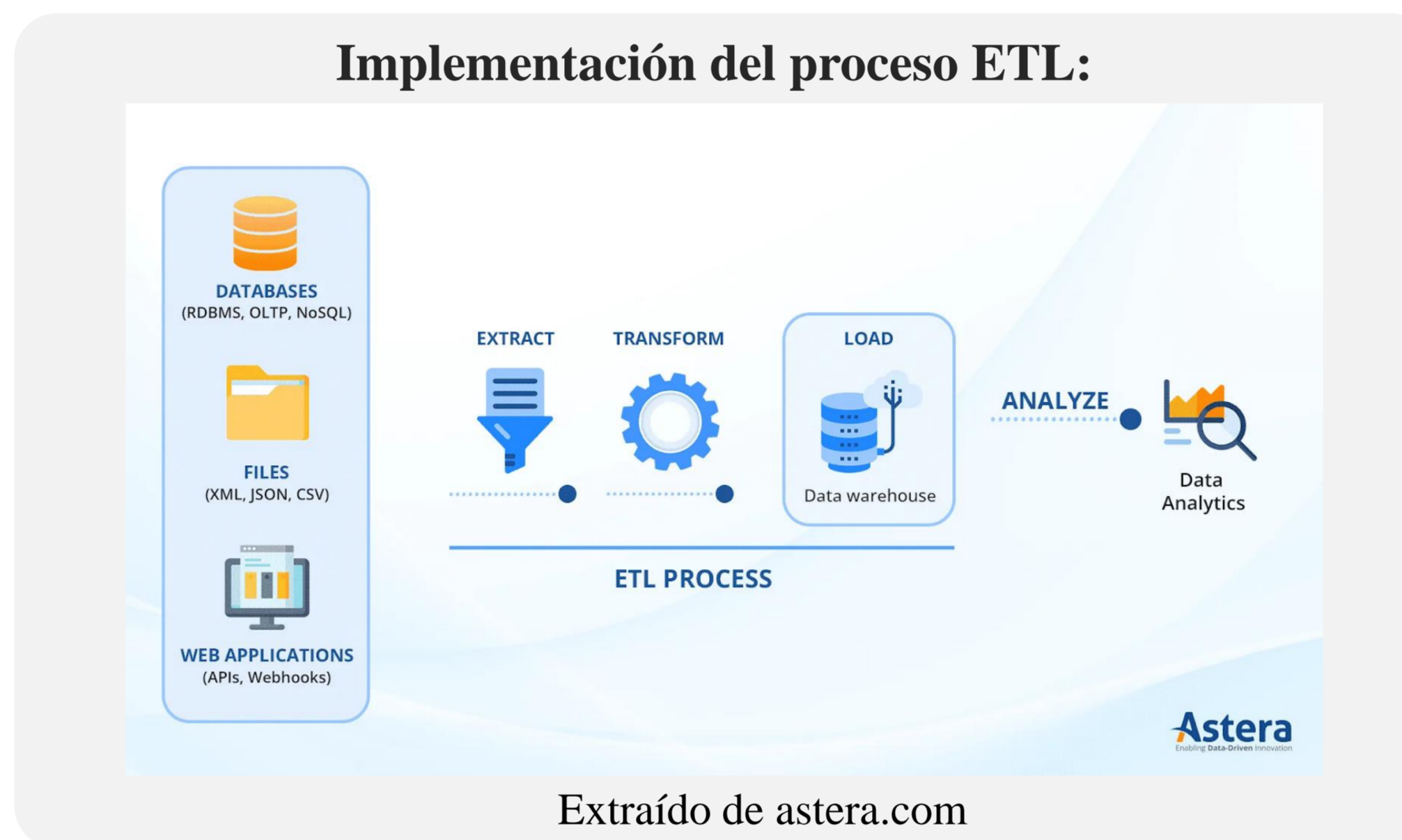
Extraído de tigobusiness.com

Metodología

Para cumplir con los objetivos planteados se realizó un análisis de los datos y sus fuentes, junto a un diseño de las nuevas tablas en la base de datos destino y la implementación de un proceso ETL. Para ello se trabajó con los lenguajes de programación Python y SQL, y el gestor de bases de datos DBeaver. Además, se usaron módulos y librerías específicas de Python como Pandas, pyodbc, cx_Oracle, re, numpy y unicodedata para establecer las conexiones necesarias a las bases de datos, extraer y transformar la información. Finalmente, se utilizó la herramienta Power BI para construir el tablero que permitiría la visualización de los datos.

Introducción

Debido al gran portafolio de productos y servicios ofrecidos por la compañía, el sector encargado de la relación comercial con otras empresas, también conocido como B2B, enfrenta un desafío muy grande con el manejo de la información, específicamente de los productos Cloud ofrecidos. Esto debido a que los datos se encuentran generalizados o mal clasificados y dispersos en distintas bases, dificultando su distinción, comprensión y análisis para la toma de decisiones.



Extraído de astera.com

Objetivos

Unificar la información del producto Cloud 360 de Tigo Business y optimizar su gestión.

- Identificar las múltiples fuentes de datos a utilizar.
- Buscar y extraer de las fuentes los identificadores únicos que correspondan a servicios Tigo Cloud y clasificarlos.
- Realizar cruces para obtener el detalle de los clientes y la facturación clasificada por tipo de servicio.
- Transformar los datos para garantizar que se ajusten a la necesidad de la solución.
- Cargar los datos transformados en nuevas tablas destino.
- Construir un tablero que presente información relevante para el análisis y la toma de decisiones.

Resultados

- Identificación y exploración de las fuentes de datos:** se identificaron cuatro tablas en distintas bases de datos que contenían información variada sobre los servicios Tigo Cloud. Algunas de las tablas tenían desde 800.000 hasta 3.000.000 registros cada una.
- Extracción de la información:** se obtuvieron aproximadamente 23,000 identificadores únicos durante la extracción, que posteriormente fueron cruzados con las bases de datos correspondientes para obtener la facturación y el detalle de los clientes asociados a estos servicios.

- Transformación de los datos:**

ABC Departamento	ABC Departamento
BOGOTA	BOGOTA D.C.
BOGOTA	BOGOTA D.C.
BOGOTA	BOGOTA D.C.
BOGOTÁ D.C.	BOGOTA D.C.
BOGOTÁ D.C.	BOGOTA D.C.
BOGOTÁ D.C.	BOGOTA D.C.

Fig 1. Estandarización de los datos

Index	PEDIDO_ID	CARACTERISTICA_ID	VALOR
0		5	
1		39	
2		2967	5/30/2024 11:43:40 AM
3		4388	None
4		1	
5		170	12
6		3682	20240531
7		4024	Cloud Server
8		4026	1
9		4027	180330
10		4028	180330

Fig 2. Tabla inicial de consulta Cloud Server

Index	PERIODO	REGIONAL	IDENTIFICADOR	NIT	F_CREACION	F_FIN	DESCRIPCION	CANTIDAD	FACTURADOR	VIGENCIA
0	202401				2024-01-30	2024-03-...	Cloud 360	1	None	12
1	202402				2024-02-10	2024-06-...	Cloud 360	1	None	12
2	202402				2024-02-12	2024-03-...	Cloud 360	1	None	36
3	202402				2024-02-13	2024-04-...	Cloud 360	1	None	12
4	202402				2024-02-14	2024-07-...	Cloud 360	1	None	24
5	202402				2024-02-22	2024-08-...	Cloud 360	1	None	12
6	202402				2024-02-22	2024-09-...	Cloud 360	1	None	12
7	202402				2024-02-23	2024-06-...	Cloud 360	1		4
8	202402				2024-02-26	2024-08-...	Cloud Server	1		22
9	202403				2024-03-05	2024-08-...	Cloud 360	1		1
10	202403				2024-03-08	2024-07-...	Cloud 360	1	None	12

Fig 3. Tabla resultante tras la transformación

VALOR_UNITARIO	VALOR_UNITARIO	F_INICIO	F_INICIO	F_FIN	TIPO	F_ESTIMADA_FIN
608.112	608112	2024-03-13 00:00:00	2024-03-13	NaT	CLOUD360	2024-02-25
992.008	902008	2024-06-11 00:00:00	2024-06-11	NaT	CLOUD360	2024-09-16
2258734	2.25873e+06	2024-03-27 00:00:00	2024-03-27	NaT	CLOUD360	2024-11-13
9552334	9.55233e+06	2024-04-16 00:00:00	2024-04-16	NaT	CLOUD360	2025-02-07
1258916	1.25892e+06	2024-07-30 00:00:00	2024-07-30	NaT	CLOUD360	2025-02-08
531.164	531164	2024-08-15 00:00:00	2024-08-15	NaT	CLOUD360	2025-02-13
540.080	540080	2024-03-08 00:00:00	2024-03-08	NaT	CLOUD360	2025-02-14
11981326	1.19813e+07	2024-06-04 00:00:00	2024-06-04	NaT	CLOUD360	2025-03-08
2494853.	2.49485e+06	2024-08-15 00:00:00	2024-08-15	NaT	CLOUD360	2025-04-14

Fig 4. Limpieza y conversión de campos numéricos

Fig 5. Cambio de formato en campos de fecha

Fig 6. Proceso de completar registros vacíos

- Carga de la información:**

MES ACTUAL	HISTORICO	ACTUAL	HISTORICO
PERIODO	PERIODO	PERIODO	PERIODO
CLIENTE_ID	CLIENTE_ID	CLIENTE	CLIENTE
Nombre	Nombre	CLIENTE_ID	CLIENTE_ID
Suscripcion	Suscripcion	GERENCIA	GERENCIA
atServicioSuscrito	atServicioSuscrito	DEPARTAMENTO	DEPARTAMENTO
Codigo_Servicio	Codigo_Servicio	MUNICIPIO	MUNICIPIO
Description_Servicio	Description_Servicio	DIRECCION	DIRECCION
Codigo_Dispartamento	Codigo_Dispartamento	ESTRATO	ESTRATO
Departamento	Departamento	REGIONAL	REGIONAL
Codigo_Municipio	Codigo_Municipio	PLAZA	PLAZA
Municipio	Municipio	LINEA	LINEA
Codigo_Concepto	Codigo_Concepto	SUB_LINEA	SUB_LINEA
Concepto	Concepto	PRODUCTO_ID	PRODUCTO_ID
Codigo_PlanFacturacion	Codigo_PlanFacturacion	PRODUCTO	PRODUCTO
Pla_Facturacion	Pla_Facturacion	SUBPRODUCTO	SUBPRODUCTO
CodelementoMedicion	CodelementoMedicion	IDENTIFICADOR_ID	IDENTIFICADOR_ID
Ciclo	Ciclo	AGRUPADOR_ID	AGRUPADOR_ID
Identificador	Identificador	ESTADO_SERVICIO	ESTADO_SERVICIO
Interfaz	Interfaz	PLAN_FACTURACION	PLAN_FACTURACION
LINEA	LINEA	INTERFAZ	INTERFAZ
SUB_LINEA	SUB_LINEA	PRODUCTO_H	PRODUCTO_H
REGIONAL	REGIONAL	TIPO_SERVICIO	TIPO_SERVICIO
PLAZA	PLAZA	PRODUCTO_H	PRODUCTO_H
PRODUCTO_H	PRODUCTO_H	EJECUTIVO	EJECUTIVO
CONCEPTO_AGRUPADO	CONCEPTO_AGRUPADO	SUSCRIPCION	SUSCRIPCION
GERENCIA	GERENCIA	CODIGO_HOGAR	CODIGO_HOGAR
SEG828	SEG828	EMPRESA	EMPRESA
EMPRESA	EMPRESA	EMPRESA	EMPRESA
ESTADO_RGU	ESTADO_RGU	EMPRESA	EMPRESA
125 FACT_TOTAL	125 FACT_TOTAL	NEW_LEVEL1	NEW_LEVEL1
NEW_LEVEL1	NEW_LEVEL1	NEW_LEVEL2	NEW_LEVEL2
NEW_LEVEL2	NEW_LEVEL2	NEW_LEVEL3	NEW_LEVEL3
NEW_LEVEL3	NEW_LEVEL3	NEW_LEVEL4	NEW_LEVEL4
NEW_LEVEL4	NEW_LEVEL4	FUENTE	FUENTE
FUENTE	FUENTE		

Fig 7. Tablas de facturación

Fig 8. Tablas con detalle del cliente

PERIODO	REGIONAL	IDENTIFICADOR	NIT	F_CREACION	F_FIN	DESCRIPCION	CANTIDAD	FACTURADOR	VIGENCIA
PERIODO	PERIODO	IDENTIFICADOR	NIT	F_CREACION	F_FIN	DESCRIPCION	CANTIDAD	FACTURADOR	VIGENCIA
PEDIDO_ID									
IDENTIFICADOR									
NIT									
F_CREACION									
F_INICIO									
DESCRIPCION									
CANTIDAD									
FACTURADOR									
VIGENCIA									
123 VALOR_UNITARIO									
123 VALOR_TOTAL									
FIN									
TIPO									
ESTADO									

Fig 9. Tabla final de altas y bajas de Cloud Server

- Diseño de tablero/dashboard:**

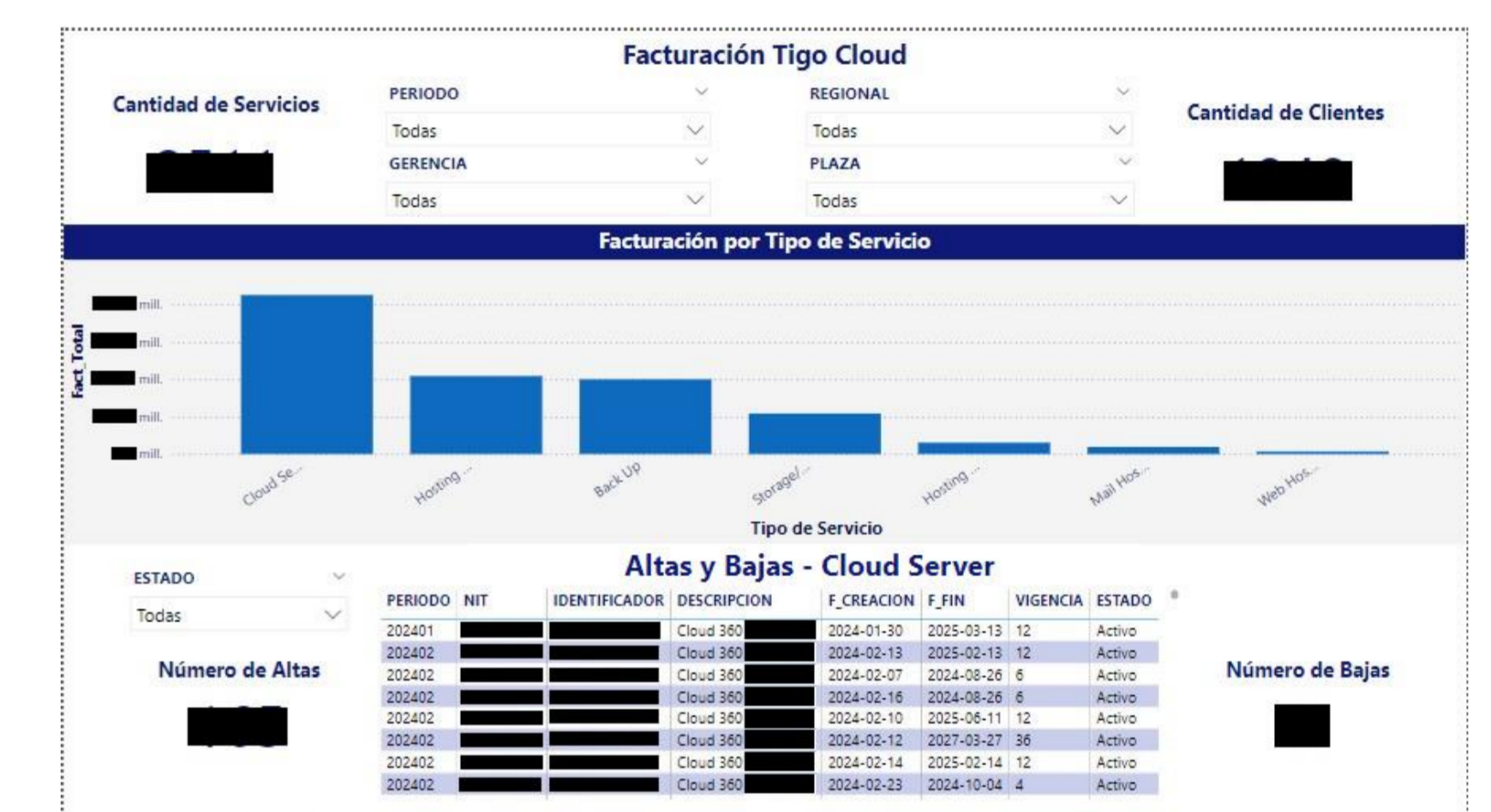


Fig 10. Tablero de facturación en Power BI

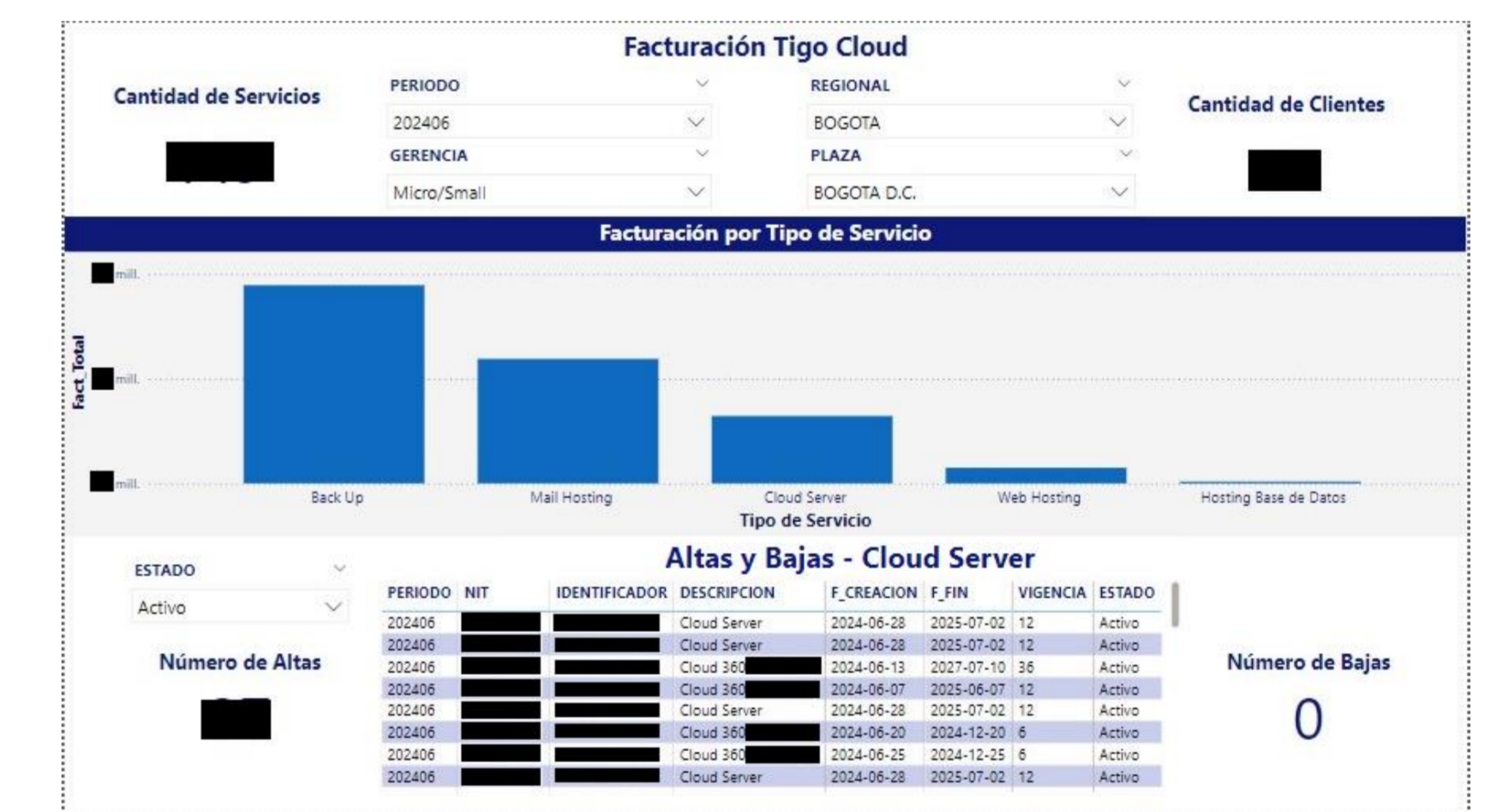


Fig 11. Variaciones en la facturación por selección de filtros

Conclusiones

- La implementación del proceso ETL permitió limpiar, estandarizar, categorizar y unificar los datos de los servicios Cloud 360 que provenían de distintas fuentes, con la finalidad de mejorar su calidad y permitir una distinción clara entre los tipos de servicios.
- La creación de las nuevas tablas permitió consolidar y almacenar toda la información resultante del proceso en una base de datos destino de fácil acceso y gestión.
- Con la estructura y organización de las nuevas tablas se facilitó la construcción del tablero en Power BI, con el cual se logró una visión clara de los datos más importantes sobre los servicios, su facturación y los clientes.
- La capacidad de filtrar los datos por período, gerencia o región en el tablero favoreció la realización de comparaciones para un análisis más detallado de la evolución de la facturación a lo largo del tiempo. Todo esto con la finalidad de tener un conocimiento auténtico de los servicios para una toma de decisiones mucho más informada.