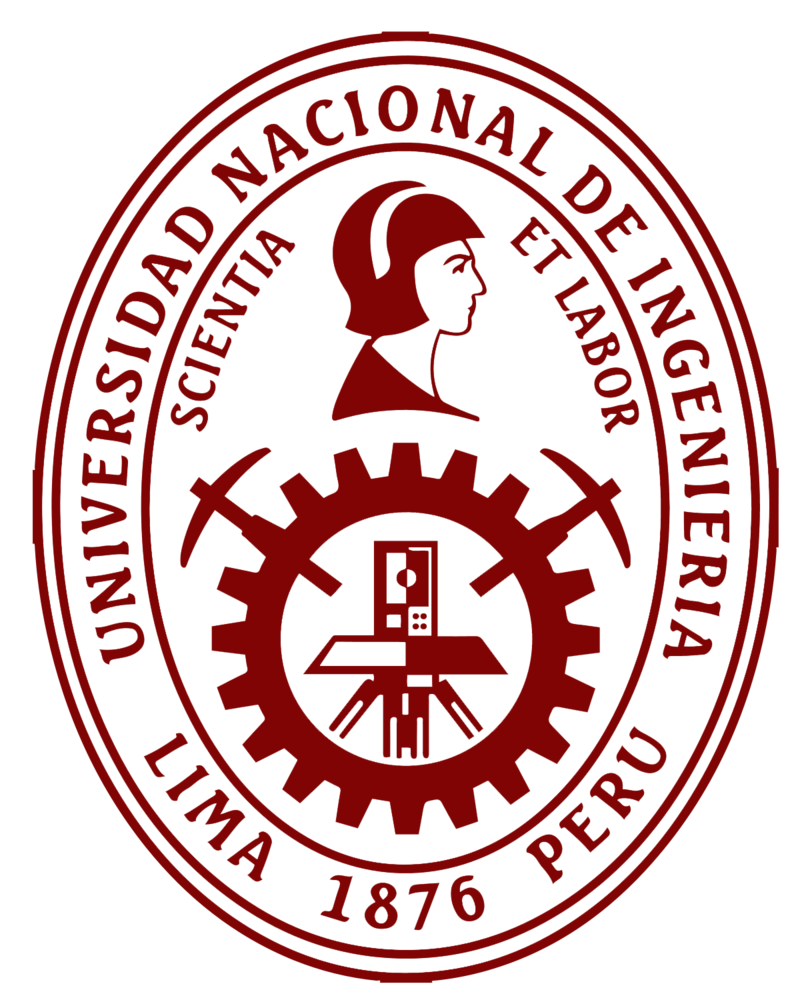
**Universidad Nacional de Ingeniería**

**Facultad de Ingeniería Industrial y de Sistemas**

**Escuela de Ingeniería de Sistemas**

**SI 807 Sistema de Inteligencia de Negocios “U”**

**Practica calificada 2**

**Caso de Estudio: “Herramientas de análisis para la SOT en el área de instalación y mantenimiento”**

**Integrantes:**

* Callupe Pardo Yoselyn Patricia 20190267H
* Garro Oré, Willian Jesús 20172115E
* Nuñez-Poma-Robert Gianpierro Jesus 20202084E

Lima – Perú

**2025**

1. **MÉTODO DE HEFESTO**
   1. **Fase 1: Análisis de requerimientos y definición del caso de negocio**
      1. **Descripción del caso de la empresa**
         1. **Sector, tamaño, procesos clave de la empresa.**

Claro Perú, que opera bajo la razón social América Móvil Perú S.A.C., forma parte de América Móvil, uno de los conglomerados de telecomunicaciones más grandes del mundo. Por su tamaño y alcance, Claro es un actor clave en el mercado peruano, ofreciendo un portafolio integral que incluye telefonía móvil y fija, internet de banda ancha y servicios de televisión por suscripción.

Nuestro caso de estudio se centra en el Área de instalación y mantenimiento residencial, un proceso operativo crítico que impacta directamente en la calidad del servicio y la satisfacción del cliente. Este proceso clave involucra una secuencia de subprocesos y actores, particularmente los contratistas externos:

* **Generación de la orden de servicio (SOT)**: El proceso inicia tras una venta exitosa o la recepción de un reclamo por falla técnica. La SOT se genera y contiene los detalles del servicio requerido (instalación o mantenimiento) y la ubicación del cliente.
* **Gestión y asignación de recursos**: Esta etapa implica la llamada al cliente para acordar una fecha y franja horaria para la visita técnica. Posteriormente, se realiza la programación y asignación de la SOT al contratista más adecuado, considerando factores como la ubicación geográfica y la disponibilidad.
* **Ejecución y cierre técnico**: El contratista realiza el trabajo en el domicilio del cliente. Una vez finalizado, se registra el cierre técnico de la SOT en el sistema, indicando el resultado (exitosa, pendiente o cancelada).
* **Control y post-servicio**: Finalmente, el área de Claro realiza el control de calidad, gestiona la liquidación con el contratista y da seguimiento a posibles reclamos posteriores.
  + - 1. **Problemas identificados**

El principal problema es la falta de control y visibilidad analítica en el proceso de Gestión de SOT y contratistas, esto genera ineficiencia y afecta al cliente:

* **Dificultad para medir el rendimiento**: La empresa no tiene una forma fácil ni rápida de saber si los contratistas están cumpliendo los plazos. Esto significa que no se puede evaluar objetivamente quién trabaja bien y quién está causando demoras.
* **Datos sucios e inconsistentes**: La información que viene de los sistemas está a menudo mal escrita o incompleta. Esta "data sucia" hace que los reportes sean poco confiables e imposibilita un seguimiento histórico correcto.
* **Demoras en el servicio**: Al no poder identificar y corregir los cuellos de botella a tiempo, las órdenes de instalación y mantenimiento se retrasan, provocando reclamos y afectando la satisfacción del cliente.
  + - 1. **Necesidades de información para la toma de decisiones.**

Para abordar los problemas descritos, el área operativa requiere información analítica específica que permita una gestión proactiva:

* **Indicadores de Desempeño (KPIs) de Contratistas**: Se necesita información para calcular métricas de rendimiento, tales como el tiempo promedio de cierre de SOT por contratista y la tasa de cumplimiento de órdenes completadas a tiempo.
* **Monitoreo del ciclo de vida de la SOT**: Es fundamental contar con datos limpios y estructurados que permitan realizar un seguimiento completo del ciclo de vida de la SOT, desde su generación hasta su cierre, identificando dónde ocurren las demoras.
* **Análisis geográfico y de servicio**: Se requiere analizar la información por dimensiones clave como el distrito de atención y el tipo de servicio para entender las variaciones en la eficiencia.
* **Estandarización de datos**: Antes de cualquier análisis, es imperativo que las herramientas de BI permitan la limpieza, estandarización y homologación de los datos fuente, como los nombres de contratistas y los números de SOT, para asegurar la calidad de la información.
  + 1. **Metodología Hefesto: Análisis de requerimientos**
       1. **Identificación de preguntas clave que el sistema de BI debe responder.**

Para resolver las ineficiencias operativas y mejorar la satisfacción del cliente, el sistema de BI debe dar respuesta a las siguientes preguntas fundamentales, derivadas de los problemas de negocio identificados:

| **Área** | **Rol de Usuario** | **Pregunta de Negocio** | **Nivel de Prioridad** | **Fuente de Datos Actual** |
| --- | --- | --- | --- | --- |
| Eficiencia del Proceso | Jefatura de Operaciones | ¿Cuál es la tasa de cancelación de SOT debido a demoras en la instalación y cómo impacta en los ingresos? | Alta | Sistema de Ventas (CRM), Sistema de Gestión de SOT. |
| Eficiencia del Proceso | Supervisor Zonal / Jefe de Operaciones | ¿Cuáles son los principales "cuellos de botella" que prolongan los tiempos de espera entre la generación de una SOT y la instalación efectiva? | Alta | Sistema de Gestión de SOT (timestamps de estados). |
| Eficiencia del Proceso | Jefatura de Operaciones | ¿Cuál es la tasa de conversión de solicitudes que se convierten efectivamente en SOT completadas y cómo varía por región o tipo de servicio? | Alta | Sistema de Ventas (CRM), Sistema de Gestión de SOT. |
| Eficiencia del Proceso | Jefe de Contratistas / Supervisor Zonal | ¿Cuál es el tiempo promedio de atención (dilación) por cada tipo de servicio y por cada equipo de trabajo o contratista? | Media | Sistema de Gestión de SOT. |
| Calidad y Rechazos | Analista de Calidad | ¿Cuáles son los motivos de rechazo más recurrentes en las SOT (por tipo de servicio, ubicación, condiciones técnicas, etc.)? | Alta | Sistema de Gestión de SOT (logs y comentarios de técnicos). |
| Calidad y Rechazos | Supervisor Zonal | ¿En qué zonas geográficas se concentra la mayor cantidad de rechazos de SOT? | Media | Sistema de Gestión de SOT. |
| Calidad y Rechazos | Analista de Calidad | ¿Cuál es la tasa de resolución en primera visita y cuáles son las principales causas para requerir una segunda visita? | Alta | Sistema de Gestión de SOT. |
| Rendimiento de Contratistas | Jefe de Contratistas | ¿Qué contratistas presentan los mejores y peores tiempos de atención y tasas de rechazo? | Alta | Sistema de Gestión de SOT. |
| Rendimiento de Contratistas | Jefe de Contratistas / Analista de Calidad | ¿Cómo se correlaciona el desempeño de un contratista (tiempo, calidad) con el nivel de satisfacción del cliente final? | Alta | Sistema de Gestión de SOT, Sistema de Encuestas (NPS). |
| Rendimiento de Contratistas | Jefatura de Operaciones | ¿Es posible predecir la efectividad futura de las contratas y la probabilidad de rechazo en ciertas zonas basándose en datos históricos? | Media | Sistema de Gestión de SOT (históricos). |

* + - 1. **Identificación de KPI's clave.**

Para medir el desempeño del proceso y responder a las preguntas clave, se han identificado los siguientes Indicadores Clave de Desempeño (KPIs), basados en los KPIs iniciales propuestos:

| **Nombre del KPI** | **Descripción** | **Fórmula** | **Unidad de medida** | **Frecuencia** | **Nivel de decisión** |
| --- | --- | --- | --- | --- | --- |
| Tasa de conversión de SOT | Mide el porcentaje de órdenes de trabajo completadas exitosamente respecto al total de generadas. | (Nº de SOT completadas / Nº total de SOT generadas) \* 100 | Porcentaje (%) | Mensual | Estratégico |
| Tasa de cancelación de SOT | Indica el porcentaje de SOT que son canceladas por el cliente antes de la instalación. | (Nº de SOT canceladas / Nº total de SOT generadas) \* 100 | Porcentaje (%) | Mensual | Estratégico |
| Tasa de rechazo de SOT | Mide el porcentaje de SOT que son rechazadas por motivos técnicos o de otra índole. | (Nº de SOT rechazadas / Nº total de SOT generadas) \* 100 | Porcentaje (%) | Semanal | Táctico |
| Tiempo promedio de Atención (Dilación) | Calcula el tiempo promedio desde la generación de la SOT hasta su cierre exitoso. | Suma de (Fecha Cierre - Fecha Creación) / Nº total de SOT completadas | Días / Horas | Semanal | Táctico |
| Índice de resolución en Primera Visita (First Visit Resolution) | Porcentaje de SOT que se completan con éxito en la primera visita del técnico. | (Nº de SOT resueltas en 1ra visita / Nº total de SOT completadas) \* 100 | Porcentaje (%) | Mensual | Operativo |

* + - 1. **Perspectivas de análisis.**

El sistema de BI deberá proporcionar vistas y dashboards adaptados a las necesidades de los diferentes roles dentro de la organización, permitiendo una toma de decisiones informada y ágil.

* **Nivel Estratégico (Jefatura de Operaciones):**
  + Necesidad de información: Analizar el costo asociado a las demoras y cancelaciones por tipo de servicio y región para evaluar la viabilidad de nuevos servicios y reestructurar procesos operativos.
  + Vista en el Dashboard: Visualizará KPIs como la Tasa de Conversión de SOT y la Tasa de Cancelación. Podrá comparar el rendimiento histórico entre regiones y contratistas para tomar decisiones a largo plazo sobre la asignación de recursos y la estrategia de expansión.
* **Nivel Táctico (Jefe de Contratistas / Supervisor Zonal):**
  + Necesidad de información: Conocer la dilación promedio por equipo de trabajo e identificar los rechazos más comunes y su ubicación geográfica para mejorar la programación de técnicos y la asignación de SOT.
  + Vista en el Dashboard: Se enfocará en la Tasa de Rechazo y el Tiempo Promedio de Atención. Utilizará filtros por zona, contratista y motivo de rechazo para identificar patrones y optimizar la asignación de SOT, así como para implementar planes de capacitación específicos.
* **Nivel Operativo (Analista de Monitoreo / Programador):**
  + Necesidad de información: Consultar en tiempo real el estado de las SOT, la disponibilidad de los técnicos y acceder a un listado de los motivos de cancelación más frecuentes para priorizar órdenes críticas y reasignar recursos ante imprevistos.
  + Vista en el Dashboard: Tendrá una vista en tiempo real del flujo de SOT, con alertas sobre órdenes que exceden el tiempo promedio de atención. Accederá a reportes detallados del Índice de Resolución en Primera Visita para tomar acciones correctivas inmediatas.
  1. **Fase 2: Análisis de Sistemas OLTP y Fuentes de Datos**

1. **Determinación de indicadores relevantes:**

* Para Sot: SOT generadas, conversión de sot, sot rechazadas, dilación de la sot.
* Para contrata: Cantidad contrata total, asignación contrata promedio, dilación promedio de la contrata.

1. **Establecimiento de correspondencia**

* El sistema de ventas de la fija se vincula con los KPI’s de las sot
* Catálogo de contrata se cruza con el sistema de venta de la fija**.**

1. **Definición de nivel de granularidad**

El nivel de granularidad de los datos es clave para garantizar que los análisis sean útiles y accionables.

* SOT: Análisis diario para identificar tendencias y patrones de trabajo de la SOT, reporte mensuales para evaluar el desempeño comercial y planificar estrategias.
* Contrata: Análisis mensual para la toma de decisiones frente a las contratas, reporte mensual para el desempeño de trabajo de las contratas.

1. **Modelo conceptual ampliado**

El DataMart de Instalaciones de Claro se estructura en torno a un modelo conceptual que integra todas las fuentes de datos y los KPI’s clave.

* DataMart de Instalaciones

Centraliza la venta de la SOT, trabajos y el ERP, incluye la información de las transacciones de la SOT, niveles trabajo y dimensiones (tipo atención, dilación, tipo rechazo)

* Dashboard Interactivo

Visualizar KPI’s clave como SOT generada en todas las regiones, cantidad de sot rechazadas, promedio de dilación por SOT, Clusterización de las SOT atendidas; esto permite a los usuarios explorar los datos en tiempo real y generar insigths accionables.

* Análisis Predictivo

Utiliza machine learning para predecir patrones de conversión, efectividad de trabajo, efectividad predictiva de contratas, rechazos futuros en ciertas zonas.

* Automatización de reportes

Generación de reportes diarios, semanales y mensuales de forma automática; adición de en tiempo real para indicadores críticos.

1. **Definición de Fuentes de Datos para el Dashboard**

**5.1. Conexión de las fuentes de datos (OLTP) al DataMart y su uso en el Dashboard**

La integración de sistemas de procesamiento de transacciones en línea (OLTP) al Data Mart (DM) es esencial para consolidar datos que faciliten el análisis y la toma de decisiones. Este proceso se lleva a cabo mediante procedimientos de Extracción, Transformación y Carga (ETL), que garantizan la calidad y disponibilidad de la información en el DW.

Identificación de las Fuentes de Datos OLTP

Para el caso de Claro, se han identificado las siguientes fuentes de datos clave:

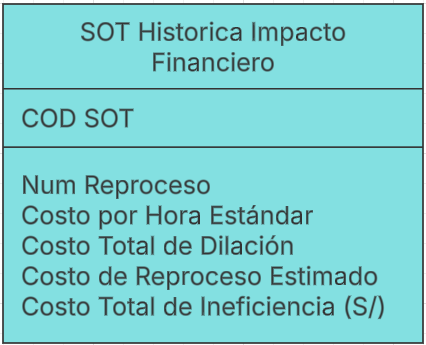
### 1. **Sistema de Gestión de SOT (OSS/BSS)**

* **Naturaleza de la fuente:** Sistema transaccional (OLTP) que administra las órdenes de servicio (instalación y mantenimiento) del negocio Fija.
* **Datos disponibles:** Código de SOT, estado de la orden (generada, en atención, cerrada, rechazada, cancelada), timestamps de cambios de estado, detalle del servicio (instalación / mantenimiento), dirección del cliente, contratista asignado.
* **Periodicidad de actualización:** En línea (transaccional) con cargas batch al D-1 hacia ODS/DataMart.
* **Uso en el Dashboard:** Medir KPIs operativos (SOT generadas, tasa de conversión, tasa de rechazo, dilación promedio, resolución en primera visita).



**2. Sistema de impacto financiero del SOT**

* **Naturaleza de la fuente**: Datos derivados del Sistema de Gestión de Órdenes de Servicio (SOT) y parámetros de costo fijos definidos por la gerencia.
* **Datos disponibles**: COD SOT, Tiempo Total de Dilación (Horas), Número de Reprocesos (Medida), Costo por Hora Estándar (S/), Costo Fijo de Traslado por Visita (S/), Costo Total de Dilación Estimado (S/), Costo de Reproceso Estimado (S/), Costo Total de Ineficiencia (S/).
* **Periodicidad de actualización**: Carga Batch al D-1 o Diaria, después de que la SOT ha alcanzado un estado de cierre.
* **Uso en el Dashboard**: Medir KPIs Financieros de Ineficiencia (Costo Total de Ineficiencia) para evaluar el impacto económico de la dilación y el reproceso por contratista, ubicación y tipo de rechazo.



3. **Sistema de Gestión de SOT (OSS/BSS)**

* **Naturaleza de la fuente:** Sistema transaccional (OLTP) que administra las órdenes de servicio (instalación y mantenimiento) del negocio Fija. Es la fuente principal de donde se originan los datos para medir la eficiencia en campo.
* **Datos disponibles:** Código de SOT, Timestamps de cambios de estado, Detalle del servicio, Dirección del cliente, Contratista y cuadrilla asignados, Logs de visitas técnicas, Datos de costos asociados, Costo Operativo Total.
* **Periodicidad de actualización:** En línea (transaccional) con cargas batch programadas (D-1) hacia el ODS (Operational Data Store) y posteriormente al Data Mart de Eficiencia Operativa.
* **Uso en el Dashboard:** Medir KPIs de eficiencia operativa y calidad del servicio para la toma de decisiones táctica y estratégica.



**5.2. Proceso de Integración al DataMart**

La integración de estas fuentes de datos al DW se realiza mediante procesos ETL, que constan de las siguientes etapas:

1. **Extracción**: Recolección de datos desde los sistemas OLTP, asegurando conexiones seguras y eficientes para minimizar el impacto en los sistemas operacionales.
2. **Transformación**: Limpieza y transformación de los datos extraídos para asegurar su calidad e integridad. Esto incluye estandarización de formatos, eliminación de duplicados, manejo de valores nulos y aplicación de reglas de negocio específicas.
3. **Carga**: Los datos transformados se cargan en el DM, donde se organizan en esquemas adecuados para el análisis, como el esquema estrella o copo de nieve.

**Frecuencia de Actualización**

La frecuencia de actualización de los datos en el dashboard depende de las necesidades de la organización y de la naturaleza de los datos:

**Naturaleza de los procesos operativos**

* El área de instalaciones y mantenimiento maneja un **alto volumen de transacciones diarias** (miles de SOT generadas, atendidas, rechazadas, pendientes o fraude).
* No es crítico analizarlas en tiempo real, sino **disponer de la foto consolidada del día anterior** para evaluar desempeño, cumplimiento de SLA y detectar cuellos de botella.

**Balance entre costo y beneficio**

* Mantener la información en tiempo real implicaría replicar todo el **OLTP (OSS/BSS, CRM, ERP)** en un **DW streaming** (Kafka, Spark Streaming, etc.), lo cual genera un alto costo de infraestructura y complejidad de gobierno de datos.
* La actualización al D-1 permite **controlar costos**, manteniendo suficiente frescura de datos para la toma de decisiones tácticas y operativas.

**Integridad y consistencia de datos**

* Muchos registros de SOT pasan por varios estados (atendido, pendiente, fraude, rechazada).
* Si se extraen en línea, se corre el riesgo de capturar **información incompleta o en transición**.
* El **corte al cierre del día** asegura que cada SOT tenga un estado final válido, facilitando la limpieza y homologación de la información en los procesos ETL.

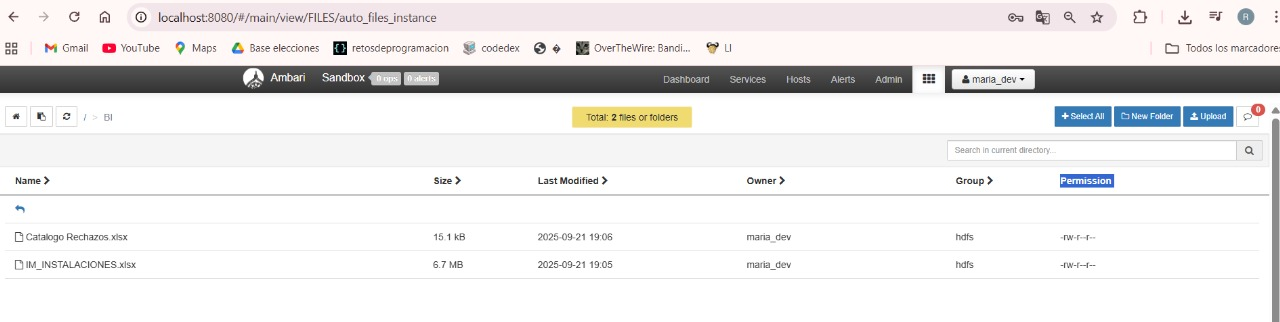
**Volumen de datos**

* Se estima que Claro puede generar **~1 millón de registros por día** en este proceso.
* Procesarlos al D-1 mediante cargas batch es más manejable que mantener un pipeline en tiempo real, donde los volúmenes podrían saturar la arquitectura de ingestión.

**Uso analítico esperado**

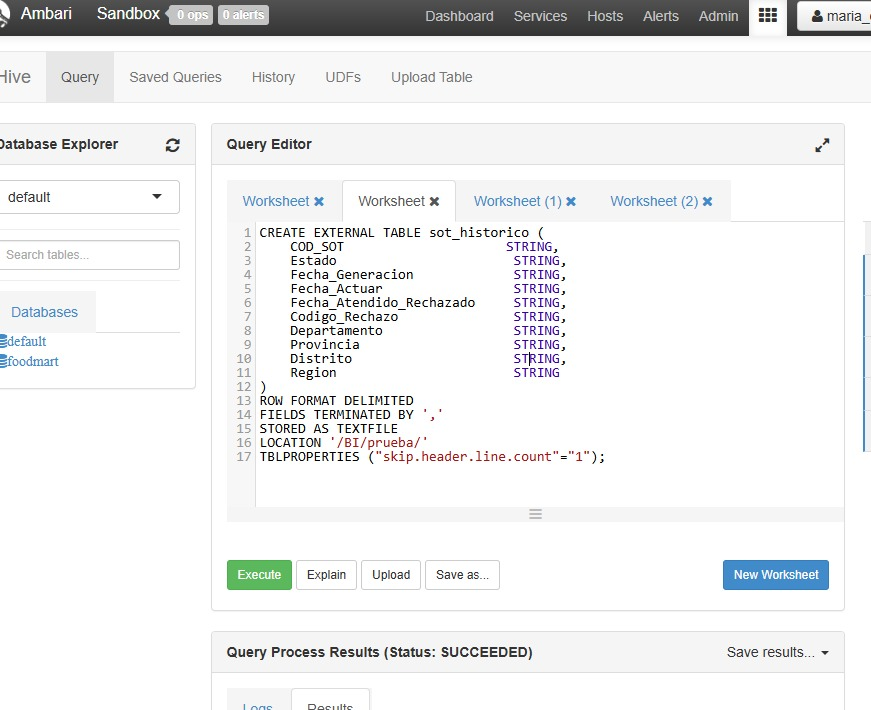
1. Los indicadores identificados (tasa de conversión, rechazo, dilación, resolución en primera visita) **no requieren decisiones en segundos**, sino monitoreo diario, semanal y mensual.
2. La vista al D-1 permite suficiente capacidad de reacción para reprogramar contratistas, detectar contratistas con bajo rendimiento y anticipar fallas operativas.
3. **EVIDENCIA TÉCNICA**
   1. **Explorando HDFS**

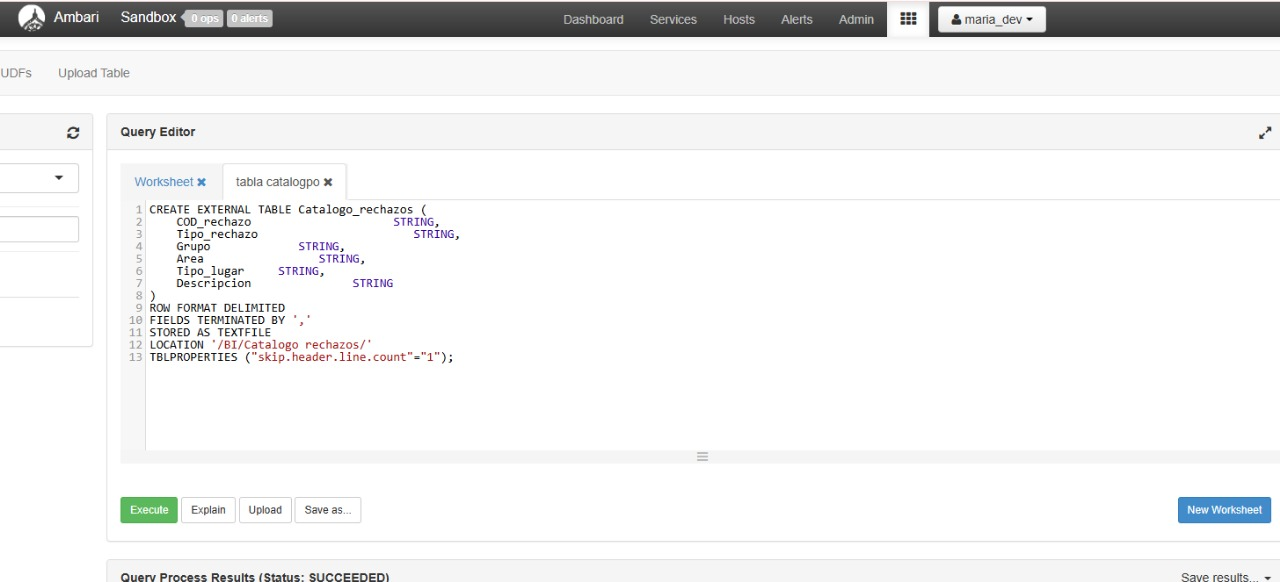
HDFS es el sistema de archivos distribuido de Hadoop, diseñado para almacenar grandes volúmenes de datos de forma fiable y eficiente en clusters de nodos. Su ventaja respecto a otros sistemas de almacenamiento tradicionales es su capacidad para dividir archivos grandes en bloques, los cuales se distribuyen a través de múltiples nodos en el clúster.



* 1. Conexión Hive con HDFS

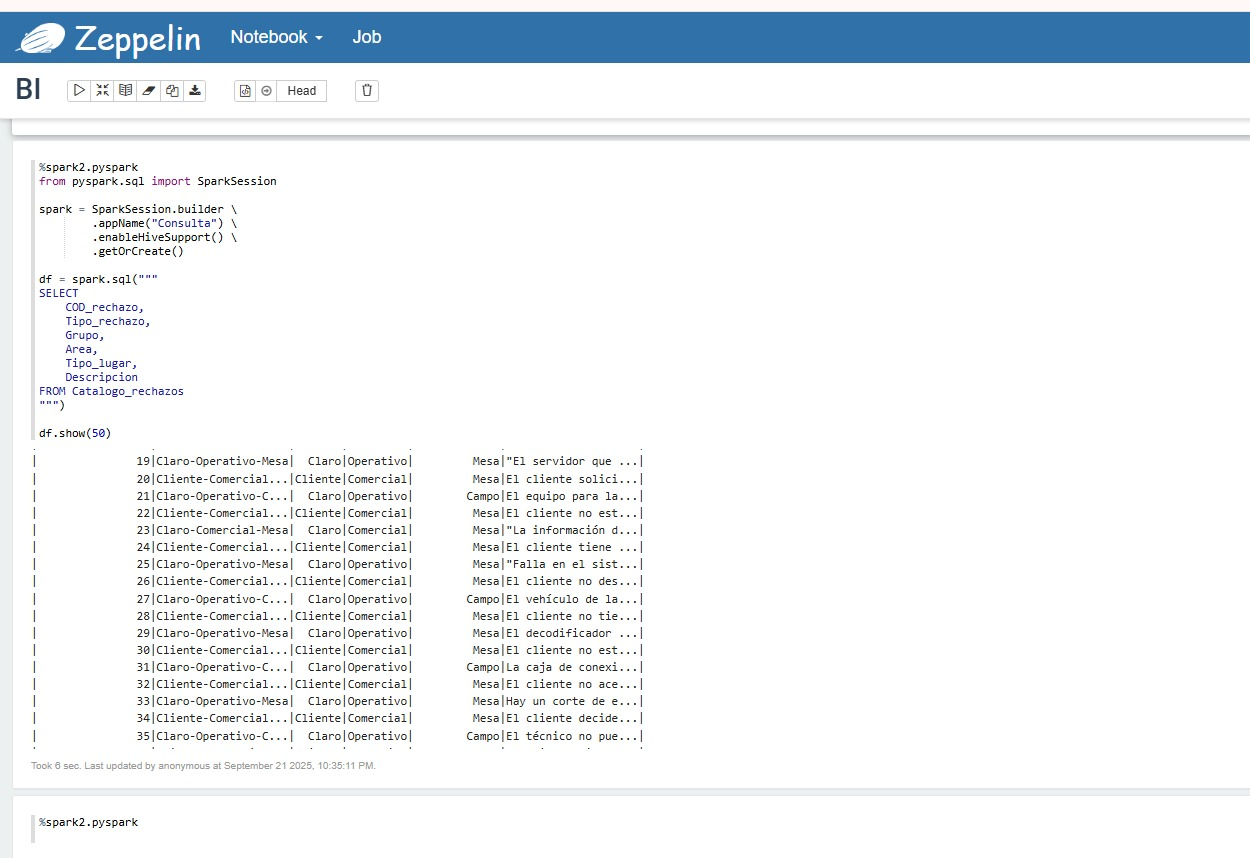
La configuración por defecto tiene límites que permiten que el sandbox funcione correctamente con recursos limitados, y cambiar estos parámetros puede desbalancear el sistema y hacer que el sandbox deje de funcionar adecuadamente.





* 1. Conectar a Datos en Hive

También podemos conectarnos con hive gracias una tabla externa. Para ello utilizamos este script.



El código proporcionado puede ser ejecutado directamente desde Apache Zeppelin usando el intérprete de PySpark (%spark2.pyspark) para interactuar con los datos en Hive.

