**PROCESO BI PARA LA EMPRESA WIN**

**Objetivo**

El propósito de este documento es detallar el proceso ETL implementado para el proyecto WIN, el cual se encarga de gestionar y analizar los datos relacionados con la red de la empresa WIN, optimizando la toma de decisiones a través de la construcción de un Data Warehouse.

Los requerimientos son los siguientes

Quiero conocer cantidad total de problemas por tipo en un tiempo determinado

Indicador: cantidad total de problemas6

Perspectiva: tipoDeProblema , tiempo

Quiero conocer el porcentaje de calidad de señal en un tiempo determinado

Indicador: Porcentaje total de calidad 2.4ghz y 5ghz

Perspectiva: tipoDeSeñal, tiempo

Quiero conocer cantidad total de usuarios que usan repetidores en un tiempo determinado

Indicador: cantidad total de usuarios

perspectiva: tiempo

Quiero conocer la suma total de repetidores en un tiempo determinado

Indicador: SumaTotalRepetidores

Perspectiva: tiempo

**Alcance**

El proceso ETL abarca la extracción de datos de la subred WIN, la transformación de estos datos en un entorno de staging, y la carga final en un Data Warehouse optimizado para análisis y generación de informes.

**Visión General**

El proyecto incluye:

* Creación de una base de datos staging (Staging\_win).
* Transformación y creación de dimensiones y hechos en el Data Warehouse (DATAWAREHOUSE\_WIN).
* Procedimientos almacenados para el cálculo de métricas específicas.

**Estructura del sistema**

**Dimensiones:**

* DimIssueType: Almacena los tipos de problemas reportados.
* DimSignalType: Almacena los tipos de señales (2.4 GHz, 5 GHz).
* DimCustomer: Información de clientes y la cantidad de repetidores utilizados.
* DimCalendar: Dimensión temporal detallada.

**Hechos:**

* FactIssue: Relaciona los problemas reportados con las fechas.
* FactSignalUse: Almacena el porcentaje de uso de señales.
* FactUserRepeaterCount: Cantidad de usuarios con repetidores.
* FactRepeaterUsage: Total de puntos de acceso en uso.

**Descripción del Proceso ETL**

**Extracción de Datos**

La extracción de datos se realiza desde la tabla DataWin, filtrando solo los registros pertenecientes a la subred WIN.

**Transformación de Datos**

La transformación incluye la limpieza, filtrado y estructuración de los datos extraídos. Se crean varias tablas de transformación para la preparación de datos antes de la carga en el Data Warehouse.

**Transformaciones Específicas**

**DimCalendar:** Se utiliza para convertir fechas en un formato analítico.

**DimIssueType:** Se pobla con valores únicos de tipos de problemas.

**DimSignalType:** Se define manualmente para almacenar los tipos de señales (2.4 GHz, 5 GHz).

**Cálculo de Tasas de Uso:** Uso de procedimientos almacenados para calcular y almacenar las tasas de uso de las señales en la tabla TransformUsageRate.

**Carga de Datos**

Los datos transformados se cargan en el Data Warehouse. Esto incluye la población de las tablas de hechos con relaciones a las dimensiones correspondientes.

**Procedimientos Almacenados**

Se implementan procedimientos almacenados para automatizar cálculos y la inserción de datos en las tablas de hechos.

**Procedimientos Detallados**

**CalculateUsageRates**

Calcula las tasas de uso de señales para una fecha específica y las almacena en la tabla TransformUsageRate.

**InsertDateUsageRate**

Itera sobre las fechas disponibles en Transform1\_win para calcular y almacenar las tasas de uso.

**5. Validación y Control de Calidad**

**5.1. Verificación de Fechas**

Se ejecutan consultas de verificación para asegurar la integridad y precisión de las fechas en las tablas de hechos.

**5.2. Consultas de Prueba**

Consultas de ejemplo para verificar la correcta integración de los datos.

**5.3. Control de Calidad**

Medidas implementadas para asegurar la calidad de los datos en cada etapa del proceso ETL.

**6. Conclusión**

El proceso ETL descrito proporciona un enfoque estructurado y eficiente para la gestión de los datos relacionados con la red de la empresa WIN, optimizando la toma de decisiones a través de un Data Warehouse bien diseñado y un proceso ETL automatizado.

las respuesta que debe responde son:

Quiero conocer cantidad total de problemas por tipo en un tiempo determinado -> indicador-> cantidad total de problemas perspectiva-> tipoDeProblema , tiempo

Quiero conocer el porcentaje de calidad de señal en un tiempo determinado-> indicador-> Porcentaje total de calidad 2.4ghz y 5ghz perspectiva-> tipoDeSeñal, tiempo

Quiero conocer cantidad total de usuarios que usan repetidores en un tiempo determinado -> indicador-> cantidad total de usuarios perspectiva -> usuarios, tiempo

Quiero conocer la suma total de repetidores en un tiempo determinado -> indicador -> SumaTotalRepetidores perspectiva-> tiempo

Usaremos los campos

Gateway\_sn: Es la id único del cliente

subnet\_name : Tipo de servicio

Issue type: Problemas que presenta el usuario

Level of Service : Calidad del servicio

Networking type: Nos muestra si solo posee router o también tiene repetidores

TOTAL APS: Cantidad de punto de acceso (REPETIDORES