

Informe de Diagnóstico de Calidad de Datos

1. Introducción

Este informe presenta un diagnóstico de calidad de datos aplicado a una base de tickets de soporte. El objetivo es identificar problemas que puedan afectar la confiabilidad de la información y proponer mejoras.

2. Dimensiones evaluadas

Dimensión	Descripción	Resultado / Hallazgos
Compleitud	Verifica la presencia de datos faltantes (nulos).	Se detectaron 3 valores nulos en campos críticos como `email_cliente`, `categoria` y `descripcion`.
Unicidad	Verifica duplicados en IDs o registros.	2 registros presentan el mismo `id_ticket`, lo que compromete la unicidad del identificador.
Validez	Evalúa si los valores respetan el formato esperado.	El campo `email_cliente` tiene 2 casos con formato inválido. Además, 1 fecha está mal escrita.
Consistencia	Revisa que los datos no se contradigan entre sí.	Existen registros con fechas de cierre (`fecha_cierre`) anteriores a la de creación.
Actualidad	Evalúa si la información está actualizada o es obsoleta.	Algunos tickets no se actualizan desde hace más de 2 años, perdiendo vigencia.

3. Propuestas de mejora

- Imputación o limpieza de valores nulos, especialmente en campos como `email_cliente` y `categoria`, esenciales para segmentación y contacto.
- Revisión de claves primarias duplicadas (`id_ticket`) y corrección manual o automatizada mediante lógica de deduplicación.
- Validación de formatos de entrada, como expresiones regulares para correos electrónicos y fechas estándar.
- Controles de integridad entre campos: asegurar que `fecha_cierre` siempre sea posterior o igual a `fecha_creacion`.
- Depuración periódica de registros inactivos o muy antiguos para mantener la vigencia del sistema.

4. Conclusión

La base presenta varios problemas que, aunque corregibles, podrían afectar decisiones operacionales o el análisis si no se atienden. Se recomienda implementar reglas de validación al ingreso de datos y realizar mantenimientos periódicos.