

Descripción del Contenido

El reporte de preparación de los datos tiene como finalidad documentar de manera clara, reproducible y estructurada todas las transformaciones aplicadas al conjunto de información antes de su uso en etapas de modelado o análisis avanzado. Este documento describe el proceso mediante el cual los datos brutos se convierten en un dataset utilizable, depurado y consistente, detallando cada paso del pipeline aplicado. Incluye la selección de atributos relevantes, la limpieza de registros, el tratamiento de valores faltantes o atípicos, la normalización de estructuras, y cualquier técnica de feature engineering que contribuya a mejorar la calidad o utilidad del dataset final.

El reporte expone los criterios utilizados para excluir, transformar o reestructurar información, así como las decisiones tomadas para asegurar que los datos cumplan con estándares de coherencia, completitud y compatibilidad con los métodos analíticos que se utilizarán posteriormente. También presenta la integración de fuentes, el formateo final del dataset y las verificaciones necesarias para garantizar que las transformaciones sean reproducibles y trazables dentro del flujo de trabajo.

En conjunto, el documento permite comprender de forma transparente cómo se construyó el dataset definitivo, ofreciendo una base sólida para validar la integridad del proceso y asegurar que los resultados del modelado se fundamenten en datos preparados de manera metodológicamente sólida. Este reporte constituye una pieza clave para el control de calidad, la auditoría del pipeline y la continuidad del proyecto en etapas de mantenimiento o iteración futura.