

Manual de Diccionario de Datos v2.0

Activos de Información: Agro-Geo & Energy-Graph

Curso: Análisis de Datos - Series de Tiempo Avanzado
Docente: Jorge Iván Padilla-Buriticá
Universidad EAFIT - Periodo 2026-1

Visión Multidimensional del Consultor

En esta fase, el análisis evoluciona de la serie temporal univariada hacia una estructura de **Datos Panel** con dimensiones espaciales y relacionales. El consultor debe evaluar no solo cuándo ocurre un evento, sino dónde (**Geo**) y cómo se propaga a través de la infraestructura (**Grafos**).

1 Dataset 1: Agroindustria y Clima (agro_*.csv)

Monitoreo de precisión en cultivos de exportación. Integra sensores mesh distribuidos geográficamente.

Variable	Descripción	Naturaleza / Propósito
Agro_1-3	Variables Hídricas	Correlación alta. $I(0)$. Humedad, Evapotransp. y RH.
Agro_4	Radiación PAR	Cíclica. Ciclo día/noche.
Agro_5-7	Índices Bióticos	No Estacionarias. $I(1)$. NDVI y Biomasa.
Agro_8-10	Suelo y Viento	Estacionarias. Ruido blanco con media constante.
Latitude	Eje Y Espacial	Posicionamiento en el Oriente Antioqueño.
Longitude	Eje X Espacial	Coordenadas decimales para clustering geoespacial.
Source_Node	Sensor ID	Nodo de origen en la topología de red mesh del cultivo.
Target_Node	Gateway ID	Nodo de destino (concentrador de datos).

2 Dataset 2: Energía y Economía (ener_*.csv)

Simulación de la Red Eléctrica Nacional. Los datos reflejan la topología de despacho y la interconexión regional.

Variable	Descripción	Naturaleza / Propósito
Ener_1-3	Mercado Spot	Correlación alta entre Demanda, Precio y Temperatura.
Ener_4	Gen. Eólica	Cíclica compleja (Intermitencia estocástica).
Ener_5-7	Factores Macro	No Estacionarias. Costo de Gas y Emisiones CO2.
Ener_8-10	Calidad Potencia	Estacionarias. Frecuencia, Voltaje y Factor de Potencia.
Latitude	Ubicación Nodo	Coordenadas de subestaciones a nivel nacional (Colombia).
Longitude	Ubicación Nodo	Permite análisis de flujos de potencia regionales.
Source_Node	Subestación	Nodo de generación o transformación primaria.
Target_Node	Nodo de Carga	Punto de consumo o bus de distribución.

3 Fundamentos de Contaminación y SNR

El análisis de señales ruidosas es el pilar de este taller. La diferencia entre los archivos `clean` y `noise` radica en la inyección de **AWGN (Additive White Gaussian Noise)**.

3.1 Definición de SNR (Signal-to-Noise Ratio)

La relación señal-ruido se define matemáticamente en decibelios (dB) como:

$$SNR_{dB} = 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right) = 10 \log_{10} \left(\frac{\sigma_{signal}^2}{\sigma_{noise}^2} \right) \quad (1)$$

Un SNR de **5dB** indica que la señal es solo ligeramente más fuerte que el ruido, dificultando la detección de tendencias. Un SNR de **15dB** representa una señal más limpia, común en sensores industriales de alta calidad.

3.2 Naturaleza del Ruido

- **Ruido Temporal:** Sigue una distribución $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$, lo que lo hace incorrelacionado en el tiempo (Ruido Blanco).
- **Jitter Geoespacial:** En las versiones `noise`, las coordenadas presentan un error aleatorio que simula fallas de precisión GPS o interferencia electromagnética en la telemetría.

Reto del Taller

¿Es posible recuperar la topología original del grafo usando los datos ruidosos? El estudiante deberá aplicar filtros de suavizado espacial y temporal antes de reconstruir las relaciones de adyacencia entre nodos.