

Diplomado en Análisis de Información Geoespacial

Tipos de datos

Autor:
M. en G. Alberto Porras Velázquez

Variables, unidades de muestreo y tipos de datos

En general, es posible hablar de tres tipos de datos:

1. Datos organizados en una secuencia a lo largo de un continuo espacial o temporal. Es posible conceptualizarlos como datos unidimensionales.
2. Datos cuyo muestreo es relativo a un contexto geográfico o espacial. Suele tratarse de datos en dos dimensiones: mapas geológicos, diagramas de contorno, análisis de superficies de tendencia, etc.
3. Datos multivariados. En cada muestra se miden dos o más atributos.

Los atributos particulares que se miden al reunir los datos se llaman variables. Estos datos son colectados en cada unidad de muestreo, que puede ser un individuo (una roca) o un elemento definido (un metro cuadrado de superficie o alguna localidad específica).

Tipos de escalas

- **Escala de relación.** Describe una variable cuyos valores numéricos indican (verdaderamente) la cantidad que se mide.
 - Existe un punto cero verdadero bajo el cual no puede existir ningún dato. Por ejemplo, el tamaño de un cristal de roca.
 - Un incremento por la misma cantidad numérica indica una cantidad igual a lo largo del rango de medidas. Ejemplo: el diámetro de un cristal de 0.2 mm y de uno de 2mm crecerán en la misma cantidad si su tamaño se incrementa 10mm.

- Una relación particular se puede mantener a lo largo de la variable. Un grano de 200 micrometros es 20 veces más grande que uno de 10 micrometros.
- **Escala de intervalo.** Describe una variable que puede ser menor que cero.
 - El cero es arbitrario. En la escala de los grados Celsius, el cero se encuentra en el punto de congelación del agua; en contraste, los grados Kelvin pertenecen a la categoría de relación de escala, en la que existe un cero absoluto (equivalente -273.15°C).
 - Un incremento por la misma cantidad (2°C) implica que ésta será constante en todo el intervalo.
 - Debido a que el cero es arbitrario, una razón particular no se mantiene a lo largo de todo el rango de la variable. Por ejemplo, la razón 6°C comparada con 1°C no es la misma para 60°C comparada con 10°C . Dicho de otro modo, es incorrecto afirmar que en ambos casos la relación es 6 a 1. En términos de la escala Kelvin estas relaciones son: $279:274\text{K}$ y $333:283\text{K}$.
- **Escala ordinal.** Se aplica cuando los valores de los datos están clasificados de acuerdo a un orden relativo.
 - Por ejemplo, cinco montañas con elevaciones de 10,000m, 4500m, 4300m, 4000m y 3984m han sido medidas con una escala de relación.
 - Si se clasifican de la más alta a la más baja, asignándoles los números 5, 4, 3, 2 y 1, los datos habrán sido reducidos a una escala ordinal.
 - Esta escala no es informativa, pues no significa que la quinta montaña sea cinco veces más alta que la número uno.

- **Escala nominal.** Se aplica a los datos clasificados de acuerdo con un atributo.
 - Ejemplos: estado civil, tipo de suelo, etc.

Datos discretos y continuos

A excepción de la escala nominal, los datos pueden ser discretos o continuos.

Sin embargo, como la escala nominal trata de atributos, sólo se puede tener el caso discreto.

- **Datos continuos.** Pueden tener cualquier valor dentro del rango. Por ejemplo, cualquier valor de temperatura es posible en el rango de 10°C a 20°C (15.3°C o 17.82°C).
- **Datos discretos.** Sólo pueden tener valores numéricos fijos dentro del rango. Por ejemplo, el número de electrones en un átomo se incrementa en enteros.

Referencias

- Steve McKillup y Melinda Darby Dyar. 2010. *Geostatistics Explained. An Introductory Guide for Earth Scientists*. United Kingdom: Cambridge University Press.