



FACULTAD
DE CIENCIAS
ECONÓMICAS



Universidad
Nacional
de Córdoba

Unidad 4. Herramientas para la visualización de datos

Docente: Mauricio Morán

Septiembre - Octubre 2023



Programa tentativo de la Unidad 4:

- **Clase 1:** Introducción y conceptos fundamentales. Tipos de datos. Tipos de gráficos. Buenas prácticas para la visualización de datos. Principales paquetes Python. Introducción a Matplotlib.
- **Clase 2:** Visualización de datos numéricos: datos, mapas, gráficos, uso de glyphs, gráficos apilados. Matplotlib-Pandas y Seaborn.
- **Clase 3:** Visualización de datos no numéricos: gráficos, redes, gráficos de componente-principal, árboles, escalado multidimensional. Presentación de scikit-learn: escalado - estandarización - normalización, PCA, árboles de decisión para clasificación y regresión.
- **Clase 4:** Reportes, dashboards y otros recursos. Comunicación de datos. Business Intelligence. Storytelling. Dashboards con Plotly.
- **Tutoria:** Desarrollo de actividad práctica.

Clase 2: Visualización de datos numéricos.

Tipos de gráficos (continuación)




¡Más gráficos!

Vamos a continuar viendo algunos nuevos tipos de gráficos:

- Scatter matrix.
- Joinplots.
- Pairplots.
- Mapas de calor.
- Mapas coropléticos.
- Mapas interactivos.

Además vamos a seguir haciendo nuevos gráficos de los tipos que vimos en la Clase 1 empleando otros paquetes.

Paquetes de Python para visualización

matplotlib +  pandas

seaborn



GeoPandas



Folium

Otros paquetes para datos geográficos:

GeoPy. Facilita a los desarrolladores de Python localizar las coordenadas de direcciones, ciudades, países y puntos de referencia en todo el mundo utilizando geocodificadores de terceros y otras fuentes de datos.

GDAL/OGR. Librería fundamental para procesar formatos de datos vectoriales y ráster.

Cartopy. Esta librería permite dibujar mapas para que el análisis y la visualización de datos sea lo más intuitivo posible.



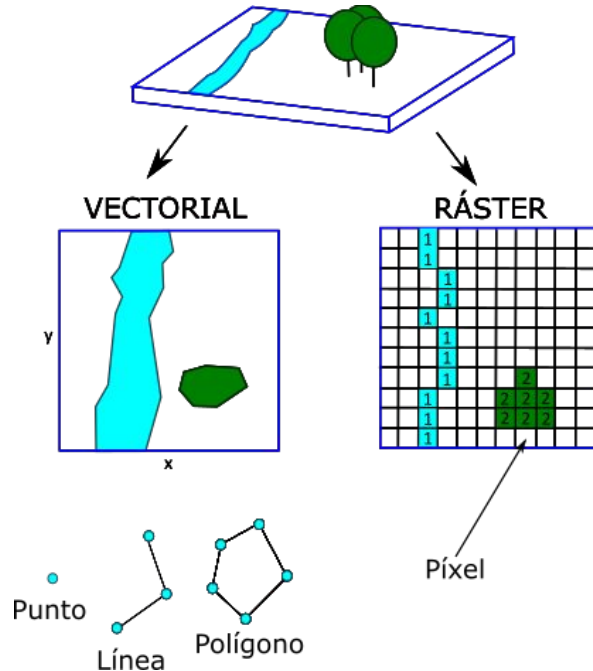
Tipos de datos espaciales

Modelo vectorial y modelo ráster

Modelo vectorial

Los datos espaciales son contruidos por geometrías y atributos. Los puntos son coordenadas x,y,z, las líneas unen puntos y los polígonos son conjuntos de líneas.

Ejemplos: shapefile, geojson, KLM.



Modelo ráster

Los datos espaciales son representados por unidades de igual tamaño, a veces llamadas píxeles. Cada celda contiene un valor para un fenómeno u observación en particular.

Ejemplos: geoTIFF, JPEG(2000).

mygisnotebook.blog



Shapefiles

Es un formato multiarchivo, los archivos mínimos son:

- .shp: almacena las entidades geométricas de los objetos.
- .shx: almacena el índice de las entidades geométricas.
- .dbf: almacena la información de los atributos de los objetos, es la base de datos.






Y algunos archivos opcionales son:

- .prj: guarda la información referida al sistema de coordenadas.
- .sbn y .sbx: almacenan el índice espacial de las entidades.
- .fbn y .fbx: almacenan el índice espacial de las entidades para los shapefiles que son inalterables.
- .ain y .aih: almacenan el índice de atributo de los campos activos en una tabla o el tema de la tabla de atributos.
- .xml: almacenan los metadatos del shapefile.

wikipedia.org/wiki/Shapefile

Shapefiles

Importante: deben estar todos los archivos en la misma carpeta / ubicación aunque sólo se importe el archivo .shp

Nombre	Tipo
 cabaxrdatos.dbf	Archivo DBF
 cabaxrdatos.prj	Archivo PRJ
 cabaxrdatos.qpj	Archivo QPJ
 cabaxrdatos.shp	Archivo SHP
 cabaxrdatos.shx	Archivo SHX

Algunas páginas para descargar archivos shapefiles:

- INDEC indec.gob.ar
- IGN - Instituto Geográfico Nacional: ign.gob.ar
- Natural Earth naturalearthdata.com
- US Census Bureau census.gov
- Eurostat ec.europa.eu

iA programar!

