

Introducción al manejo de datos espaciales usando R

Módulo 1

Sergio A. Barona
Laboratorio de Economía Aplicada
Departamento de Economía y Finanzas

Pontificia Universidad Javeriana de Cali

21-04-2025

1 Presentación del Curso

2 R como herramienta SIG

3 Paquetes en R para datos espaciales

4 Tipos de datos espaciales

5 Referencias

Presentación del Curso

Datos espaciales: la ubicación importa

Más allá de la elaboración de mapas, los datos espaciales permiten formular y responder preguntas que dependen de la **localización geográfica** y, por tanto, no se puede resolver únicamente mediante la observación directa de los datos ([Bivand, Pebesma & Gómez-Rubio, 2008](#)).

¿Qué tipo de preguntas se pueden abordar?

- ¿Cuáles son los patrones espaciales y principales factores climáticos de la transmisión de la malaria en tres zonas fronterizas de Brasil, Venezuela y Guyana? ([Wangdi et al., 2022](#)).
- ¿Cuál es la distribución espacial y cuáles son los factores predictivos de la violencia doméstica contra las mujeres en Etiopía? ([Seid et al., 2021](#)).
- ¿Influye la distancia geográfica entre el candidato y el votante en la probabilidad de voto en las elecciones parlamentarias de Reino Unido? ([Arzheimer & Evans, 2012](#)).

¿Y usando datos socioeconómicos?

- ¿Influyen los factores espaciales en la convergencia del ingreso regional en Estados Unidos? ([Rey & Mountouri, 1999](#)).
- ¿Varían los procesos que generan la pobreza infantil en diferentes subregiones de Estados Unidos? ([Curtis, Voss & Long, 2012](#)).

Rey & Mountouri (1999)

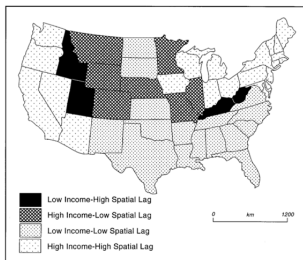


Fig. 4. Local Mean statistics per capita income, 1929

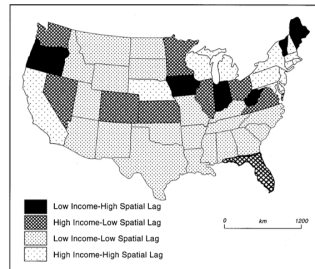


Fig. 5. Local Mean statistics per capita income, 1994

Curtis, Voss & Long (2012)

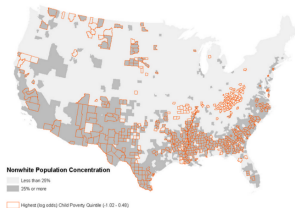


Figure 2.
Racial Concentration Regimes based on Percent Non-white and Counties with High Child Poverty in 2000.

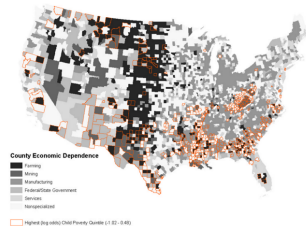


Figure 3.
Industrial Structure Regimes based on Economic Dependence and Counties with High Child Poverty in 2000.

R como herramienta SIG

¿Por qué R en general?

R tiene las siguientes ventajas:

- R es un software libre.
- R es multiplataforma.
- R tiene una **sofisticada capacidad para hacer gráficos**, incluyendo paquetes gráficos especializados.
- R tiene librerías que cubre un amplio rango de la metodologías de la estadística y las matemáticas (series temporales, optimización matemática, inferencia estadística, etc.)
- Existe una comunidad activa que ha promovido el incremento en su número de usuarios ([The R Project](#), [R Contributor Site](#), [RStudio Community](#) y [R Bloggers](#)).

¿Por qué R como herramienta SIG?

La definición de [Burrough and McDonnell \(1998, p. 11\)](#):

Un conjunto poderoso de herramientas par recopilar, almacenar, transformar y visualizar datos espaciales del mundo real para un conjunto específico de propósitos. [Es] verificar, manipular y analizar datos que están referenciados espacialmente en la Tierra

R es una buena elección para el manejo y análisis de datos espaciales, dadas sus capacidades robustas de análisis y visualización. Entre sus ventajas:

- **R base** ya incluye muchas funciones para leer, visualizar y analizar datos espaciales.
- La funcionalidad base se ha complementado con una amplia gama de paquetes para datos raster y vectoriales, metadatos geográficos, muestreo espacial, análisis espacial, geoestadística, machine learning espacial, entre otros. (Para una descripción detallada: [Bivand & Nowosad \(2025\)](#))
- Comunidades activas en este campo: [r-spatial](#) y [rspatial](#)
- Muchos paquetes utilizan clases compartidas y estándares comunes, lo cual reduce la duplicación de esfuerzos.

Paquetes en R para datos espaciales

Paquete sf

Simple features (características simples) hace referencia a un estándar formal (ISO 19125-1:2004) que describe cómo se representan los objetos del mundo real, enfocándose en su geometría espacial. También establece cómo almacenarlos en bases de datos y las operaciones geométricas que se deben definir.

- Una **característica (feature)** se entiende como una entidad u objeto del mundo real.
- Un **conjunto de características** puede conformar una única **característica** más compleja.
- Las **características** tienen una **geometría** que indica su localización y poseen **atributos** que describen otras propiedades no espaciales.



```
install.packages("sf")
```

```
library(sf)
```

Paquete terra

- terra es un paquete de R para el análisis de datos espaciales, enfocado en datos raster.
- Permite crear, leer, manipular y escribir datos raster de forma eficiente.
- Sustituye el paquete raster, con una interfaz similar, pero es más simple, rápido y potente.
- Puede trabajar con archivos raster demasiado grandes para cargar en la memoria (RAM). El paquete terra sólo carga la estructura del archivo, i.e., número de filas y columna, extensión espacial, nombre del archivo, etc.



```
install.packages("terra")  
library(terra)
```

terra: clases

El paquete terra está construido sobre varias clases de objetos. Las clases `SpatRaster` y `SpatVector` son las más importantes.

Clase	Descripción
<code>SpatRaster</code>	Representa datos raster (grillas) de múltiples capas. Este objeto almacena varios parámetros fundamentales que lo describen (número de columnas, filas, extensión espacial y sistema de referencia de coordenadas).
<code>SpatVector</code>	Representa datos vectoriales (puntos, líneas, polígonos) y sus atributos no-espaciales
<code>SpatExtent</code>	Define la extensión espacial de los objetos de la clase <code>SpatRaster</code>

Paquetes antiguos

Hasta 2016, el paquete `sp` proporcionaba clases para datos espaciales vectoriales y raster, pero sus representaciones eran anteriores a los estándares internacionales más modernos y eficientes para datos vectoriales (incluidos en `sf`).

terra es un paquete que reemplaza el paquete **raster** con una interfaz similar, pero más simple y rápido.

Otros paquetes:

- `rgdal` ([Bivand, Keitt & Rowlingson, 2023](#))
- `rgeos` ([Bivand & Rundel, 2022](#))
- `maptools` ([Bivand & Lewin-Koh, 2022](#))

Paquetes antiguos (cont.)

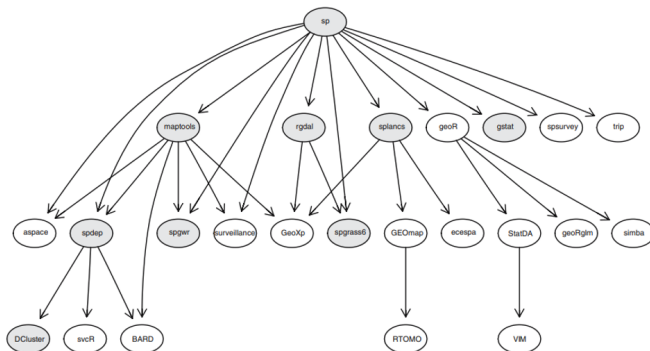
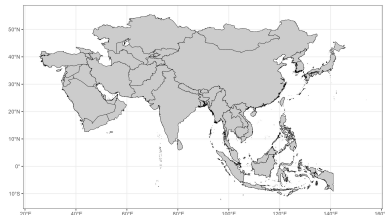
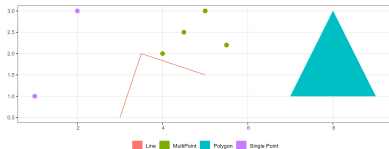


Figure 1: Paquetes de CRAN que dependen (de manera directa o indirecta) de sp. Fuente: Bivand, Pebesma & Gómez-Rubio (2013)

Tipos de datos espaciales

Datos vectoriales

Vector data (datos vectoriales) es un modelo general usado para representar entidades espaciales como puntos, líneas y polígonos, junto con sus atributos. El paquete `sf` permite trabajar con datos vectoriales.



Tipos de geometrías sf

Las geometrías más comunes son las siguientes:

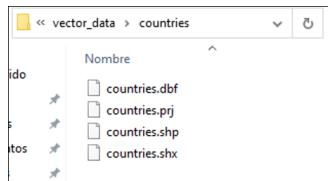
Tipo	Descripción
POINT	Punto de dimensión cero.
LINESTRING	Secuencia de puntos conectados por segmentos de línea; geometría unidimensional.
POLYGON	Geometría bidimensional con área positiva; secuencia de puntos que forman un anillo cerrado que no forma intersección con sí mismo. Un polígono consta de uno o más anillos.
MULTIPOINT	Conjunto de puntos individuales
MULTILINESTRING	Conjunto de líneas independientes
MULTIPOLYGON	Conjunto de polígonos representados en una sola geometría
GEOMETRYCOLLECTION	Conjunto de geometrías de cualquier tipo (excepto otra GEOMETRYCOLLECTION)

Otras: **CIRCULARSTRING**, **COMPOUNDCURVE**, **CURVEPOLYGON**, **ULTICURVE**, **MULTISURFACE**, **CURVE**, **SURFACE**, **POLYHEDRALSURFACE**, **TIN**, **TRIANGLE**. (Véase: [Simple Features for R](#))

Formato Shapefile

El **shapefile** es un formato común para almacenar datos vectoriales. No es un sólo archivo, sino un conjunto de archivos relacionados que deben estar en la misma carpeta. Los archivos principales son los siguientes:

- **_____.shp**: almacena la geometría (puntos, líneas, polígonos, etc.)
- **_____.shx**: índice posicional de la geometría
- **_____.dbf**: atributos asociados
- **_____.prj**: sistema de coordenadas e información de proyección
- **_____.sbn** y **_____.sbx**: índice espacial de las entidades
- **_____.xml**: metadatos espaciales en formato XML.



Nota: para trabajar correctamente con un shapefile, es imprescindible contar al menos con los archivos **.shp**, **.shx** y **.dbf** juntos en la misma carpeta con el mismo nombre base.

Leer/escribir datos vectoriales

La función `st_read()` del paquete `sf` permite leer archivos shapefile.

```
# Cargar paquete  
library(sf)
```

```
# Definir directorio de trabajo  
setwd("path")
```

```
# Definir ruta y capa  
st_read(dsn = "path2", layer = "layer")
```

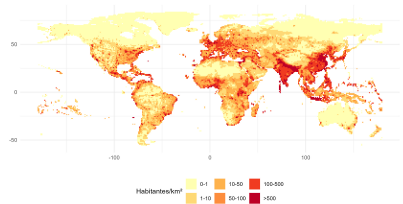
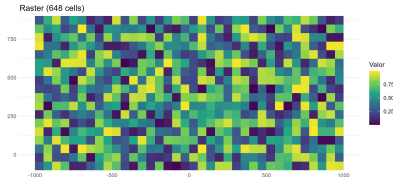
La función `st_write()` del paquete `sf` se utiliza para escribir datos espaciales a archivos:

```
st_write(obj = x, dsn = "path2", layer = "path3")
```

Siguiente paso: ejecutar el script *modulo1.R*.

Datos raster

Los **datos raster** son una estructura de datos que representan el espacio como un conjunto de celdas del mismo tamaño (o píxeles), ordenadas en filas y columnas, y que pueden almacenar uno o más valores para cada celda.



Los **datos raster** se utilizan comúnmente para representar fenómenos continuos en el espacio, tales como la elevación.

Formato GeoTIFF

Los datos raster generalmente se almacenan en formato GeoTIFF, cuya extensión es .tif.

```
# Cargar paquete  
library(terra)
```

Usando la función `rast()` del paquete `terra`, podemos crear objetos raster.

```
# Crear objeto raster  
r <- rast(nrows=20, ncols=20, # Número de celdas (dimension 20x20)  
          xmin=0, xmax=360) # Coordenadas mín y máx.
```

También podemos utilizar esta función para leer archivos con la extensión .tif.

```
# Leer raster  
terra::rast("raster_file.tif")
```

Siguiente paso: ejecutar el script *modulo1.R*.

Referencias

Referencias

- Pebesma, E. (2023). *Simple Features for R (sf) – Vignette*. <https://cran.r-project.org/web/packages/sf/vignettes/sf1.html>
- Hijmans, R. J. (2023). *Terra: Spatial Data Analysis*. CRAN package
- Arzheimer, K., & Evans, J. (2012). Geolocation and voting: Candidate–voter distance effects on party choice in the 2010 UK general election in England. *Political Geography*, 31(5), 301–310. <https://doi.org/10.1016/j.polgeo.2012.04.006>
- Bivand, R., Keitt, T., & Rowlingson, B. (2023). *Rgdal: Bindings for the Geospatial Data Abstraction Library*. CRAN package
- Bivand, R., & Lewin-Koh, N. (2022). *Maptools: Tools for Handling Spatial Objects*. CRAN package
- Bivand, R., & Rundel, C. (2022). *Rgeos: Interface to Geometry Engine – Open Source (GEOS)*. CRAN package
- Bivand, R. S., Pebesma, E., & Gómez-Rubio, V. (2008). *Applied Spatial Data Analysis with R*. Springer.
- Burrough, P. A., & McDonnell, R. A. (1998). *Principles of Geographical Information Systems*. Oxford University Press.
- Curtis, K. J., Voss, P. R., & Long, D. D. (2012). Spatial variation in poverty-generating processes: Child poverty in the United States. *Social Science Research*, 41(1), 146–159. <https://doi.org/10.1016/j.ssresearch.2011.07.007>

- Moraga, P. (2023). *Spatial Statistics for Data Science: Theory and Practice with R*. Chapman & Hall/CRC Data Science Series. ISBN: 9781032633510.
- Rey, S. J., & Montouri, B. D. (1999). US Regional Income Convergence: A Spatial Econometric Perspective. *Regional Studies*, 33(2), 143–156.
<https://doi.org/10.1080/00343409950122945>
- Seid, M. A., Melese, Y. Y., & Tolossa, T. (2021). Spatial distribution and predictors of domestic violence against women: Evidence from analysis of Ethiopian demographic health survey 2016. *BMC Women's Health*, 21, 231.
<https://doi.org/10.1186/s12905-021-01465-4>
- Wangdi, K., Lambrechts, T., & Clements, A. C. A. (2022). Spatial patterns and climate drivers of malaria in three border areas of Brazil, Venezuela and Guyana, 2016–2018. *Sci Rep*, 12, 10666. <https://doi.org/10.1038/s41598-022-14012-4>