

Análisis de datos geográficos con R.

Fundamentación.

En las últimas décadas, a través de los avances tecnológicos, la humanidad ha generado un volumen de datos enorme que crece exponencialmente a cada año. Si bien esta producción inmensa de datos permite analizar cuestiones y/o temas novedosos, también implica la necesidad de utilizar nuevas herramientas más adecuadas para manejar grandes volúmenes de datos.

El lenguaje de programación R surge en 1993 en la Universidad de Auckland con la característica de poseer una gran potencia a la hora de procesar grandes volúmenes de datos, generar gráficos y realizar cálculos estadísticos. Este lenguaje tiene la característica de ser gratuito y de código abierto. Esto implica en primer lugar una gran accesibilidad al estar exento de pago por licencia. Por otro lado, el código abierto ha permitido el desarrollo del lenguaje a través de una comunidad de usuarios que ha ido creando una innumerable cantidad de librerías que han facilitado su utilización y expandido sus posibilidades.

El análisis de datos espaciales es fundamental en numerosos campos, ya que permite entender patrones y relaciones en el espacio geográfico que de otro modo serían difíciles de detectar. Su importancia radica en la capacidad de agregar una dimensión espacial a los datos, lo cual es crucial para la toma de decisiones informadas en áreas que dependen de la ubicación y distribución espacial.

Objetivos.

A lo largo del curso se buscará que los asistentes sean capaces de buscar bases de datos de su interés para comprender su estructura y realizar las operaciones correspondientes para poder extraer información de valor a través de código en lenguaje R. Dentro de las capacidades que se espera que puedan realizar los estudiantes se encuentra:

- Comprender la lógica del lenguaje R en su dimensión espacial
- Búsqueda y carga de bases de datos espaciales en el entorno de R
- Cálculo de distancias
- Cálculo de áreas

- Generación de gráficos estáticos e interactivos
- Generar gráficos de polígonos, líneas y puntos
- Crear ruteos entre direcciones
- Conexión con WFS / WMS
- Conexión con APIs
- Geocodificación (obtención de coordenadas a partir de una dirección)
- Geocodificación Inversa (obtención de una dirección a partir de coordenadas)

Requisitos.

En relación a los requisitos, se requieren conocimientos previos básicos en R (principalmente comandos básicos de tidyverse y ggplot) y tener instalado R y RStudio. Se requiere tener computadora propia con buen funcionamiento, internet de buena conexión y un espacio libre de distracciones para poder realizar la cursada.

Modalidad de cursada.

La cursada está planificada para 4 encuentros de 2 horas reloj cada uno en modalidad a distancia a través de google meet.

Las clases quedarán grabadas y a disposición de los asistentes.

Cada clase atravesará distintas instancias:

- Presentación de los temas a trabajar en la clase
- Una sección expositiva en la que los alumnos irán escribiendo código a medida que el profesor explica las funciones y presenta el código
- Espacio de dudas y consultas sobre los temas trabajados durante la clase o en encuentros previos
- Un desafío semanal a resolver fuera del horario de clase

Cada encuentro tendrá material propio enviado a los inscriptos:

- Material explicativo de los temas de la clase
- Código ejecutado durante la clase
- Archivo de desafío semanal

Contenidos.

Unidad 1.

Introducción a la manipulación de datos espaciales con R. Lectura, visualización y escritura de bases de datos espaciales.

- Librería sf() para la manipulación de bases de datos espaciales
- Bases de datos de líneas, polígonos y puntos
- Etiquetado de polígonos en mapas
- Join de base de datos con datos espaciales
- Redes de transporte

Unidad 2.

Trabajo con datos sociodemográficos y su representación espacial.

- Cálculo de áreas
- Cálculo de densidad poblacional
- Mapas con referencia
- Pirámides poblacionales y su distribución espacial
- Mapas de calor

Unidad 3.

Geoservicios. Conexión con aplicaciones.

- Conexión con servicio WFS
- Conexión con servicios WMS
- Conexión con APIs
- Cálculo de distancias

Unidad 4.

Mapas interactivos.

- Introducción a OpenStreetMaps
- Paquete leaflet para mapas interactivos
- Mapas base (Stamen)
- Ruteo de direcciones
- Geocodificación

-Geocodificación inversa

Docente.

El curso está a cargo de Diego Pacheco. Lic. En Sociología (UBA), doctorando en Ciencias Sociales (UBA) y docente del seminario de investigación “Explorando la periferia” de la carrera de Sociología (UBA).

Trabajó como analista de datos tanto en el ámbito público como privado.

Ámbito público:

- Ministerio de Desarrollo Social de la Nación (Programa POTENCIAR TRABAJO)
- Ministerio de Seguridad de la Nación (Secretaría de políticas criminales)
- Gobierno de la Ciudad de Buenos Aires (Dirección General de Sistemas de Información Sanitaria)

Ámbito privado:

- Data engineer en proyectos de machine learning para YPF

Becas:

UBACyT (Universidad de Buenos Aires Ciencia y Técnica): Actualmente su proyecto de tesis doctoral incluye trabajo con información cuantitativa de barrios populares en el conurbano e información georreferenciada.

Ganador beca FUNDATOS II (fundación Fundar). Proyecto sobre detección de déficit habitacional en el conurbano bonaerense a partir de algoritmos de aprendizaje automático.