



# Procesamiento y Visualización de Datos Espaciales en R

Profesor: José Luis Texcalac Sangrador

Laboratorio — 04

El procesamiento de datos es una tarea fundamental que antecede al análisis de datos y el paquete `{dplyr}` es uno de los más utilizados para tal fin. Forma parte del núcleo de paquetes de `{tidyverse}` y nos permite trabajar bajo el enfoque del **Tidy Data**. Proporciona un conjunto de funciones (a menudo referidas como “verbos”) que facilitan las tareas comunes que el procesamiento de datos requiere, tales como la selección de columnas, filtrado de filas, reordenamiento y agregación.

## ¿Qué se espera de ti en este laboratorio?

Esta actividad pretende que usted desarrolle habilidades para el procesamiento de datos utilizando la función `select()` del paquete `{dplyr}`, para ello se requiere que aplique los conocimientos adquiridos en clase y que explore mediante búsquedas en internet la información disponible y de utilidad para la realización del laboratorio.

## Indicaciones por considerar:

- Suba su laboratorio a la plataforma [Google Classroom](#) a más tardar el jueves 10 de octubre a las 16:00 h.
- Lo que debe usted entregar para evaluar su laboratorio es:
  - Archivo (script) con extensión **.R**
- La cuenta de correo electrónico para el envío de laboratorios y cualquier asunto relacionado al curso es: ✉ [jtexcalac@insp.edu.mx](mailto:jtexcalac@insp.edu.mx)
- Nombre a su archivo con el siguiente patrón: **L04\_Nombre**
- Cualquier duda puede publicarla en [Google Classroom](#) y su profesor o compañeros le pueden auxiliar.
- Revise las dispositivas de este tema, ahí encontrará código que le apoye a la realización de este laboratorio

## Instrucciones

Genere un nuevo script y en él programe las siguientes tareas.

- El script debe ir con encabezado y secciones
  - Comente cada uno de sus ejercicios del script
  - Puede tomar como referencia el estilo de comentar que se muestra en los distintos scripts de apoyo disponibles en [Google Classroom](#).
  - Para este laboratorio usted requiere utilizar la función `select()` del paquete `{dplyr}`.
1. Importe el archivo “[Promedio horarios de o3.csv](#)” que usted descargó de la página del [Sistema de Monitoreo Atmosférico](#) del Gobierno de la Ciudad de México.
    - Visite nuevamente la página, diríjase al apartado de **Monitoreo** y en el menú contextual que se despliega elija **Estaciones de monitoreo**, revise la tabla, la requiere para la resolución de este ejercicio.
    - Genere una malla de datos que incluya sólo a las estaciones de la Ciudad de México, nombre al objeto como `ozono_cdmx`.
    - Genere una malla de datos que incluya sólo a las estaciones de la Ciudad de México, nombre al objeto como `ozono_edomex`.
  2. Importe la malla de datos que usted descargó del Sistema Urbano Nacional ([SUN](#)), nombre al objeto como `sun_2018` y realice las siguientes actividades .
    - Todas las mallas de datos que usted genere para este ejercicio deberán contener las primeras 5 columnas del objeto `sun_2018`.
    - Visite la página del [SUN](#) y consulte el diccionario de datos, lo requiere para resolver el ejercicio.
    - Genere una malla de datos con las columnas cuyos nombres indican que contienen información de censos de población o de proyecciones de población, nombre a su malla como `pob_sun`.
    - Genere una malla de datos con las columnas cuyos nombres indican que contienen información de encuestas intercensales del año 2015, nombre a su malla como `pob_2015`.

### Material de apoyo

Puede usted apoyarse de los siguientes recursos para la realización de su laboratorio.

- Función `select()`
- [Estadísticas descriptivas en R](#)