



Procesamiento y Visualización de Datos Espaciales en R

Profesor: José Luis Texcalac Sangrador

Laboratorio — 08

Los [factores](#) se utilizan para trabajar con variables categóricas, es decir, variables que tienen un conjunto fijo y conocido de valores posibles. Los valores que un factor puede contener están delimitados por los niveles, es decir, las diferentes categorías presentes en una variable categórica. Los niveles pueden estar codificados tanto en valores numéricos como caracteres. En el entorno de [Tidyverse](#) los factores se trabajan con el paquete [forcats](#) que nos facilita varias de las tareas básicas y primordiales para el procesamiento de este tipo de información.

¿Qué se espera de ti en este laboratorio?

Esta actividad pretende que fomentes tu habilidad en la redacción de código y sus resultados utilizando el lenguaje de marcado [Markdown](#), se busca que fortalezca el aprendizaje en el uso de “verbos” del paquete [dplyr](#) para la manipulación de datos y en el procesamiento de datos de tipo factor. Se busca que explores mediante el código de apoyo que se te proporciona y búsquedas de internet, la información disponible y de utilidad para la realización del laboratorio. Se espera que la edición que programe en su archivo [R Quarto](#) contenga el formato, diseño y calidad que usted considere adecuada para la presentación de un reporte html.

Indicaciones por considerar:

- Suba su laboratorio a la plataforma [Google Classroom](#) a más tardar antes del inicio de la próxima sesión (jueves 30 de octubre).
- Cualquier duda puede publicarla en Google Classroom y su profesor o compañeros le pueden auxiliar.
- La cuenta de correo electrónico para cualquier asunto relacionado al curso es: [✉ jtexcalac@insp.edu.mx](mailto:jtexcalac@insp.edu.mx)
- Lo que debe usted entregar para evaluar su laboratorio es:
 - Archivo comprimido [.zip](#)
 - El archivo [zip](#) deberá contener la [carpeta](#) con las salidas del laboratorio, el archivo con el código con extensión [.qmd](#) y el archivo [html](#)

- Nombre a su archivo con el siguiente patrón: **L08_Nombre**. El laboratorio NO requiere enviarse por correo electrónico, evítelo.

Instrucciones:

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

- Comente cada uno de sus ejercicios y el resultado.
- Revise los vínculos que se le proporcionan en el material de apoyo, contienen ejemplos que le auxiliarán a realizar sus ejercicios.
- Los gráficos que se le pidan deberán ir editados en color y usando algún **tema**.
- Comente el resultado que muestra cada gráfico.

1. Para esta parte del laboratorio usted utilizará el archivo **covid_factores.rds** que usted debe descargar del sitio de la clase en **Google Classroom**. La malla de datos proviene de las fuentes oficiales mexicanas en dónde cada fila representa a una muerte por COVID-19, cada columna describe las características de cada caso.

- Importe la malla de datos y nombre a su objeto como **covid**
- Convierta a factor las variables: sexo, intubado, indígena, diabetes, epoc y asma, para la variable sexo utilice las etiquetas del catálogo de variables. Guarde su objeto como **cov_def** y ejecute el **summary()**
- Seleccione las columnas: entidad_res, municipio_res, fecha_ingreso, fecha_sintomas, fecha_def, edad, sexo, intubado, indigena, diabetes, epoc, asma y guarde nuevamente su objeto como **cov_def**.
- Genere las columnas: **day**, **month**, **year** y **week** que contengan el día, mes, año y número de semana respectivamente. Guarde su objeto como **cov_def**.

2. Usando el comando **count** genere los procesamientos para resolver los siguientes ejercicios

- Tabla que muestre las muertes por entidad en orden descendente.
- Tabla que muestre las 10 entidades con mayor número de muertes del sexo femenino.
- Tabla que muestre las 10 entidades con mayor número de muertes del sexo masculino.

- iv. Revise las páginas que se le sugieren al final de este archivo, revise los contenidos que le ayudarán a realizar los siguientes ejercicios:
- v. Genere una tabla que contenga las 10 entidades con mayor número de defunciones y el porcentaje de cada una (use el comando `tabyl()`).
- vi. Genere una tabla de contingencia que nos muestre, para la Ciudad de México ("09"), las frecuencias por sexo según su condición de intubado (use el comando `tabyl()`).

Material de apoyo

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

- [The janitor package](#)
- [Un CRAN a la vez: janitor](#)
- [Tablas de frecuencia con la función tabyl\(\)](#)
- [tabyls: a tidy, fully-featured approach to counting things](#)
- [Bibliographies with R and Quarto](#)