Reporte Automatizado

Área de análisis

2023-09-06

# Introducción

R markdown es un lenguaje de texto sencillo y fácil de usar que se utiliza para combinar el código R, los resultados del análisis de datos (incluidos gráficos y tablas) y comentarios escritos en un único documento con un formato agradable y reproducible (como un informe, una publicación, un capítulo de tesis o una página web).

Realizar reportes a través de R markdown, nos permite automatizar y replicar aquellos análisis de problemas públicos en materia de seguridad y procuración de justicia que requieren actualización constante y oportuna.

El objetivo de este primer reporte automatizado es integrar diversos temas que ya hemos visto, utilizando Rmarkdown. En este sentido, reconocerás pedazos de código y terminlogía que hemos utlizado en temas anteriores. Es recomendable que ejecutes este ejercicio una vez que hayas completado todas los ejercicios de sesiones pasadas.

# A) Origen de la base de datos

En esta sección retomamos el marco conceptual del repositorio de información y documentación de metadatos.

## I. Descripción del metadato

| Etiqueta del elemento | Descripción del contenido/ Contenido esperado |
| --- | --- |
| Área responsable de la elaboración del metadato |  |
| Fecha de conclusión de la documentación |  |
| Número de identificación del documento de metadatos |  |

## II. Descripción de proyecto estadístico o base de datos

| Etiqueta del elemento | Descripción del contenido/ Contenido esperado |
| --- | --- |
| Título |  |
| Tipo de proyecto estadístico |  |
| Resumen |  |
| Unidad de análisis |  |
| Clase de datos |  |
| Instrumento(s) de captación de la información |  |
| Cobertura geográfica |  |
| Universo de estudio |  |

## III. Descripción de las variables

Nombre del archivo:

| No. | Nombre de variable | Tipo de datos | Número de registros | Valores faltantes | Porcentaje de valores faltantes | Número de valores distintos |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | fecha\_hechos | Date | 64,973 | 0 | 0.00 | 120 |
| 2 | mes\_hechos | character | 64,973 | 0 | 0.00 | 4 |
| 3 | hora\_hechos | hms:difftime | 64,973 | 0 | 0.00 | 1,431 |
| 4 | agencia | character | 64,973 | 0 | 0.00 | 116 |
| 5 | alcaldia\_hechos | character | 64,973 | 0 | 0.00 | 18 |
| 6 | municipio\_hechos | character | 653 | 64,320 | 0.99 | 92 |
| 7 | latitud | numeric | 64,973 | 0 | 0.00 | 60,657 |
| 8 | longitud | numeric | 64,973 | 0 | 0.00 | 59,800 |
| 9 | hora | integer | 64,973 | 0 | 0.00 | 24 |
| 10 | dia | character | 64,973 | 0 | 0.00 | 7 |
| 11 | ano | character | 64,973 | 0 | 0.00 | 1 |
| 12 | mes | character | 64,973 | 0 | 0.00 | 4 |

La fase inicial de Análisis Exploratorio de Datos (EDA por las siglas en inglés) es crucial para obtener mejores resultados en las fases posteriores del análisis de datos.

Las estimaciones del tiempo dedicado a EDA y a la preparación de datos oscilan entre el 60 y el 80% del tiempo total dedicado a un proyecto de Análisis de Datos.

Sin embargo, existen paquetes en R que nos permitan acelerar nuestro EDA sin dejar de hacer un trabajo exhaustivo.

Con una sola línea de código utilizando los métodos de SmartEDA, podemos obtener cada uno de los siguientes datos: \* Información mucho más rica \* Mejor formato de salida para facilitar la lectura \* Resumen exhaustivo multipunto de cada característica del conjunto de fechas \* Visualización de análisis univariantes y bivariantes

La tabla resume cada característica del conjunto de datos, incluido el tipo de variable, el % de casos que faltan y varios valores únicos para cada parte.

# B) Descripción de bases de datos

Completa los campos vacios de este reporte.

Describe la base datos con base en lo que conocemos hasta ahora

# C) Ejercicio

Identifica los elementos en este primer reporte de Rmarkdown Haz knit del documento para obtener el word Agrega un chunk con un código que genera una gráfica (no olvide agregar las librerías) Agrega el texto que describa la gráfica Cambia el tipo de documento a pdf