

Trabajo Práctico 2

Análisis Exploratorio de Datos de un Dataset Real

Objetivo: Aplicar técnicas de importación, limpieza, manipulación y visualización de datos.

Descripción:

- Se entrega conjunto de datos desordenado y "sucio" relacionado con datos climáticos de Argentina, provistos por el Servicio Meteorológico Nacional (ver <https://www.smn.gob.ar/descarga-de-datos>). Se entregan datos desde el 26-11-2017 hasta 31-12-2023.
- Los datos diarios están incluidos en archivos de texto, por ejemplo "datohorario20230102.txt", donde se indican el año, mes y día.
- Los campos incluidos son FECHA, HORA [HOA], TEMP [°C], HUM [%], PNM [hPa], DD [gr], FF [km/hr], NOMBRE. NOTA: cada línea de datos tiene una longitud fija de 100 caracteres. Cada variable ocupa un número fijo de caracteres.
- Además se entregan los siguientes archivos:
 - smn_estaciones.csv: con datos de las estaciones meteorológicas de Argentina
 - smn_precipitaciones.txt: con datos de precipitaciones
- Combinar los datos provistos en un único dataset que contenga todos los datos relevantes.
- Importar los datos y limpiarlos usando técnicas aprendidas en clase (dplyr, tidyr).
- Realizar un análisis exploratorio completo, incluyendo estadísticas descriptivas y visualizaciones con ggplot2 o plotly. Realizar gráficos georeferenciados interactivos.
- Identificar y comunicar los hallazgos más importantes del dataset.
- Identificar y manejar valores faltantes y duplicados.
- Transformar variables según sea necesario (por ejemplo, convertir fechas a un formato adecuado, categorizar variables, etc.).
- Graficar.

Entregable: informe en R Markdown, Jupyter Notebook, R estándar o Quarto Document que incluya código, visualizaciones y conclusiones.

Bonus:

- Realizar un análisis de series temporales.
- Ajustar modelos de pronóstico (se recomienda usar [Prophet](#)) y evaluar su rendimiento.
- Realizar pronósticos para los próximos 12 meses.