

11 de julio de 2024

Introducción a la Ciencia de Datos

Tarea 3 - Grupo 5

Agustín Ghazarian
Sofía Gervaz

Tarea 3

Se ha seleccionado la base de datos “climate change indicators”¹. Este conjunto de datos contiene indicadores de cambio climático para diferentes países desde 1961 hasta 2022. El dataset tiene información por país, de cambios de temperatura respecto a una referencia. Tiene además del nombre de cada país, su nomenclatura según código ISO2 e ISO3, el tipo de indicador presentado, la fuente del mismo y la unidad.

Posibles problemas en la calidad de datos incluyen:

- Valores faltantes
- Inconsistencias en los códigos de país (ISO2, ISO3)
- Errores tipográficos
- Valores duplicados en columnas de nombre o nomenclatura de país (Country, ISO2, ISO3)
- Unidades de medida inconsistentes (Unit)
- Fuentes de datos incompletas o incorrectas (Source)
- Descripciones incompletas o incorrectas (CTS_Full_Descriptor)
- Problemas de formato y codificación
- Tratamiento de países que hayan cambiado en el periodo de estudio (por ejemplo URSS)

El set de datos permite visualizar la evolución del cambio climático en los distintos países. Se podrían responder las siguientes preguntas:

- ¿Cuál es la tendencia de la temperatura global promedio en las últimas décadas?
- ¿Qué país presenta un incremento mayor?
- ¿Qué región presenta un incremento mayor?
- ¿Se observan cambios de tendencia por país? ¿Se pueden vincular estos cambios en las tendencias con la situación de desarrollo del país?
- ¿Se observa algún cambio en los años correspondientes a grandes crisis? Por ejemplo: pandemia COVID-19
- ¿Se está desacelerando el incremento en la temperatura global?
- ¿Si se sigue esta tendencia se logra alcanzar los objetivos del acuerdo de París de limitar el incremento de temperatura a 1.5°C?

El proceso a seguir incluiría las siguientes etapas:

- Análisis exploratorio de datos
- Preprocesamiento de datos (ver datos faltantes, inconsistencia en etiquetas, etc)
- Análisis de tendencias (globales, por país, por región)
- Visualizaciones para la presentación de resultados

¹ <https://www.kaggle.com/datasets/tarunrm09/climate-change-indicators>