

Tarea 4

Profesor: Leo Ferres

Valor: 4 puntos de homework

Librerías clave: geopandas, pandas, matplotlib, rasterio, rasterstats, xarray, scikit-learn, seaborn, folium, plotly

Análisis Espacio-Temporal de Contaminación Atmosférica

Objetivo:

En esta tarea, aprenderás a trabajar con datos de calidad del aire usando Python para visualizar y analizar patrones de contaminación atmosférica por regiones en un país. Crearás mapas temáticos de diferentes contaminantes, estudiarás sus variaciones temporales y espaciales, y analizarás la relación entre calidad del aire y factores geográficos, meteorológicos y antropogénicos.

Parte 1: Búsqueda de Datos

1. **País asignado:** Usa el mismo de las tareas anteriores.
2. **Datos requeridos (búscalos tú):**
 - **Divisiones administrativas** (nivel 1 o 2): usa el mismo shapefile limpio de la tarea anterior.
 - **Datos de contaminación atmosférica:** busca datos para al menos 1 año en formato netCDF, CSV o series temporales para:
 - Material particulado (PM2.5 y PM10) - OpenAQ, CAMS
 - Ozono troposférico (O3) - Sentinel-5P
 - Dióxido de nitrógeno (NO2) - Sentinel-5P
 - Dióxido de azufre (SO2) - Sentinel-5P
 - **Datos meteorológicos:** mismos de la tarea anterior
 - **Datos de población:** densidad poblacional (WorldPop, GHSL), mismos de la tarea anterior

Parte 2: Mapas de Contaminantes y Series Temporales

1. Carga el shapefile de divisiones administrativas con **GeoPandas**.
2. Para cada contaminante (PM2.5, PM10, O3, NO2, SO2):
 - Carga los datos temporales.

- Calcula promedios diarios, mensuales y anuales por unidad administrativa.
 - Crea un mapa temático para el promedio anual.
 - Identifica “hotspots” de contaminación (áreas con valores significativamente altos).
3. Para tres ciudades o regiones importantes:
 - Grafica las series temporales diarias/mensuales de los contaminantes.
 - Identifica patrones diarios, semanales y estacionales.
 - Compara los niveles con estándares internacionales (OMS, EPA).

Parte 3: Análisis de Tendencias Temporales

1. Para cada contaminante por unidad administrativa:
 - Calcula promedios mensuales y anuales para el período disponible.
 - Utiliza modelos de descomposición de series temporales para separar:
 - Tendencia a largo plazo
 - Patrón estacional
 - Componente residual
 - Crea mapas temáticos mostrando las tendencias a largo plazo.
2. Identifica:
 - Regiones donde la calidad del aire está mejorando o empeorando.
 - Patrones estacionales dominantes.

Parte 4: Episodios de Contaminación Crítica

1. Identifica episodios de alta contaminación (valores superiores al percentil 95):
 - Fechas y duración de los episodios.
 - Distribución espacial durante estos episodios.
2. Para un episodio crítico específico:
 - Crea mapas animados que muestren la evolución del episodio.
 - Correlaciona con variables meteorológicas (temperatura, precipitación, viento).
 - Investiga posibles causas (incendios forestales, emisiones industriales, etc.).

Parte 5: Correlaciones con Factores Externos

1. Investiga la relación entre contaminación atmosférica y:
 - Densidad poblacional
 - Variables meteorológicas (temperatura, precipitación, viento)
 - Día de la semana (patrones laborales vs. fin de semana)
2. Realiza análisis de correlación y visualízalo con:
 - Matrices de correlación
 - Gráficos de dispersión
 - Mapas bivariados

Parte 6: Modelado Predictivo

1. Desarrolla un modelo simple para predecir niveles de PM2.5 o NO2:
 - Usa como predictores: otros contaminantes, variables meteorológicas, día de la semana, mes
 - Prueba diferentes algoritmos (regresión lineal, random forest, etc.)
 - Evalúa el rendimiento del modelo (RMSE, R^2)
2. Visualiza la importancia de las variables en la predicción.
3. Utiliza el modelo para predecir niveles de contaminación en diferentes escenarios (ej. aumento de temperatura, reducción de emisiones).

Entrega

1. Un archivo `.ipynb` o `.py` con tu código bien comentado.
2. Un archivo Markdown (`README.md`) explicando:
 - Fuentes de datos utilizadas.
 - Metodología de procesamiento.
 - Resultados y conclusiones principales.
 - Recomendaciones para mejorar la calidad del aire en regiones críticas.
3. Incluye:
 - **Cinco mapas temáticos** (uno por cada contaminante).
 - **Tres gráficos de series temporales** para diferentes regiones.
 - **Un mapa de tendencias** para un contaminante principal.
 - **Una visualización de episodio crítico** (mapa o animación).
 - **Dos gráficos de correlación** con factores externos.
 - **Una visualización de resultados del modelo predictivo.**

Estructura de Directorios

Organiza tu proyecto con la siguiente estructura:

```
project/
|
+- data/                                # Datos
|  |
|  +- raw/                             # Datos originales sin procesar
|  +- processed/                       # Datos procesados listos para análisis
|    |
|    +- air_quality/                  # Datos de contaminantes procesados
|    +- meteo/                       # Datos meteorológicos procesados
|    +- timeseries/                  # Series temporales procesadas
|    +- geo/                         # Datos geoespaciales procesados
|
+- notebooks/                         # Jupyter notebooks
|
+- models/                            # Modelos predictivos entrenados
|
```

```

+- output/                                # Gráficos y mapas generados
| |
| +- maps/                                # Mapas temáticos
| +- timeseries/                          # Gráficos de series temporales
| +- animations/                          # Visualizaciones animadas
|
+- README.md                              # Documentación del proyecto

```

Fecha de entrega: Domingo 25, 2025

Criterios de evaluación:

Criterio	Puntos
Mapas de contaminantes correctamente generados	0.75
Análisis de series temporales de calidad del aire	0.75
Identificación y análisis de episodios críticos	0.75
Análisis de correlaciones con factores externos	0.75
Modelado predictivo e interpretación	0.75
Calidad de visualizaciones y análisis crítico	0.25

Recursos Recomendados:

- Para datos de contaminación atmosférica:
 - OpenAQ - API para datos de estaciones de monitoreo
 - Google Earth Engine Sentinel-5P - datos satelitales
 - CAMS - Copernicus Atmosphere Monitoring Service
- Para metadatos y estándares de calidad del aire:
 - OMS Air Quality Guidelines
 - US EPA NAAQS
- Tutoriales:
 - Working with Air Quality Data in Python (long), Data Analysis with Python | Air Quality (shorter)
 - Visualizing Air Pollution with Folium
 - (Time Series Analysis with Prophet) -> no estrictamente necesario