

PARTE 1

1) Sitios de datos abiertos y tipos de datos:

- datos.gob.ar: Portal de datos abiertos de Argentina que ofrece datos sobre temas variados como salud, educación, transporte y medio ambiente, recogidos de distintas entidades gubernamentales.
- data.gov: Portal de datos abiertos de Estados Unidos con datasets sobre economía, agricultura, clima, ciencia y tecnología, entre otros.
- data.europa.eu: Plataforma europea de datos abiertos, proporciona información sobre áreas como empleo, finanzas, seguridad y energía, además de datos comparativos entre países de la Unión Europea.

2) Diferencias entre datos públicos y datos abiertos:

Los **datos públicos** son aquellos generados o bajo custodia de una entidad gubernamental y están disponibles para el público, aunque su acceso o uso puede estar restringido. Ej: estadísticas censales publicadas por una agencia gubernamental que solo se pueden ver pero no descargar.

Los **datos abiertos** pueden ser de origen público o privado, y se caracterizan por estar accesibles, reutilizables y compartibles sin restricciones, siempre bajo licencias específicas. Ej: un dataset de calidad del aire descargable en formato .csv con licencia que permite su uso y distribución.

8 características que deben cumplir los datos **públicos** para ser **abiertos**:

- Completos: disponibles y sin restricciones para su apertura.
- Primarios: recolectados en la fuente de origen, lo + específico posible, no en forma agregada ni modificada.
- Oportunos: disponibles tan rápidos como sea necesario para preservar su valor.
- Accesibles: disponibles para el rango más amplio de usuarios y propósitos.
- Procesables: estructurados de forma tal que permitan un procesamiento automático.
- No discriminatorios: disponibles para cualquier ciudadano, sin requerir un registro.
- No propietarios: disponibles en un formato sobre el cual ninguna empresa de software tenga un control exclusivo.
- Licencia de uso libre: deben tener una licencia que especifique que se trata de datos abiertos, y que defina bajo qué condiciones estos datos pueden ser reutilizados.

3) Tipos de licencias en datos abiertos:

- PDDL (Open Data Commons Public Domain Dedication and License): permite el uso, reutilización y adaptación sin restricciones.
- ODC-By (Open Data Commons Attribution License): permite la reutilización siempre y cuando se haga referencia a la autoría o fuente de los datos.
- ODbL (Open Data Commons Open Database License): permite la reutilización siempre que se reconozca la autoría de la información original y se mantenga la misma licencia en las obras derivadas, de las cuales se puede restringir su uso si, además, se distribuye una versión sin dichas restricciones.

- Creative Commons Atribución 4.0 Internacional: posibilita copiar y redistribuir el material en cualquier medio o formato, además de adaptar y transformarlo para cualquier propósito, incluso comercial. Tiene 4 condiciones: atribución, no comercial, sin obras derivadas, compartir igual.

4) Análisis posibles con un dataset de contaminación del aire:

- ✓ Análisis temporal: ver variaciones de la calidad del aire en distintas estaciones o años para detectar patrones.
- ✓ Análisis geoespacial: comparar niveles de contaminación entre ciudades para identificar áreas con mayor riesgo.
- ✓ Análisis de impacto: relacionar niveles de contaminación con indicadores de salud pública o eventos meteorológicos.
- ✓ Pronósticos: utilizar series temporales para predecir niveles futuros de contaminación en función de variables meteorológicas u otras condiciones.

OTRA:

Se van a poder hacer análisis para comparar los niveles de contaminación en las ciudades, ver los niveles en cada tipo de contaminación, impacto socioeconómico y analizar las tendencias temporales de los niveles en las ciudades.