

¿Qué es un conjunto de datos?

Una serie de datos vinculados entre sí y agrupados dentro de un mismo sistema de información para su potencial reutilización. Los conjuntos de datos o *datasets* son accesibles a través del portal de datos abiertos.

¿Qué son los datos abiertos?

Datos abiertos (open data en inglés) es una iniciativa global, ligada a las políticas de gobierno abierto, que persigue que los datos y la información, especialmente los que posee el sector público, se publiquen de forma abierta, regular y reutilizable para todo el mundo, sin restricciones de acceso, copyright-derechos de autor, patentes u otros mecanismos de control.

Los datos abiertos pueden ser reutilizados y redistribuidos libremente por cualquiera, con la única condición de atribución a su fuente o reconocimiento de su autoría.

Los datos abiertos gozan de la capacidad de ser interoperables, esto es, la capacidad de vincular -o entremezclar- diferentes conjuntos de datos de diversos sistemas y organizaciones, de manera que pueden ser localizados entre las diferentes fuentes de distintos organismos.

La interoperabilidad consiste en establecer una estructura de la información estándar para poder relacionar el contenido y los elementos de los diferentes conjuntos de datos, facilitando así la explotación de diferentes conjuntos de datos independientemente de su procedencia. Esta capacidad de componer y conectar componentes es esencial para construir sistemas grandes y complejos.

Consideraciones del conjunto de datos

Los datos de penetración son para provincias, no para comunidades autónomas.

En el conjunto de datos hay parámetros e indicadores:

- Parámetros:

Son elementos de un sistema que permite clasificarlo y poder evaluar algunas de sus características.

En este conjunto de datos los parámetros son: Operador, segmento, tecnología de acceso, provincia, año, etc...

- Indicadores:

Son las variables cuantitativas (discretas o continuas) que determinan el valor de una serie de parámetros asociados en el universo muestral.

Unos ejemplos de indicadores son: Líneas, accesos, penetraciones, etc...

Los datos están tabulados, estructurados y con las siguientes características:

- Fila única de cabecera.
- Registro único por fila.
- Estructura de datos vertical.
- No contiene totales ni agrupaciones.
- Campo de fechas codificado siguiendo el estándar ISO-8601.
- Cada una de las columnas incluye un parámetro que caracteriza el indicador final.
- Si ese dato no dispone de algún parámetro, se presenta como un valor "N/A". (ejemplo: Las líneas de Banda ancha fija mayorista, en la columna de "Segmento" hay un valor "N/A", ya que ese dato no dispone de segmento).
- En cada fila puede tener valor en varios indicadores, como es el caso de las líneas y la penetración, pero representan el mismo registro. Ya que el indicador penetración es un cálculo de las líneas con la población.

Diccionario de datos

Columna	Descripción
Año	Año
Provincia	Provincia
CCAA	Comunidad autónoma
Servicio	Parámetro de Servicio (Telefonía fija, telefonía móvil, ...)
Concepto	Concepto (Líneas, accesos, ...)
Operador	Operador
Tecnología de Acceso BAF	Parámetro de Tecnología de banda ancha fija minorista (xDSL, FTTH, HFC...)
Tecnología de Acceso	Parámetro de tecnología asociada a los accesos instalado (FTTH, DOCSIS 3.0,...)
Tipo de Estaciones Base	Parámetro de tipo de estación base (LTE, 5G, ...)
Tipo de BA Mayorista	Parámetro de Tipo de tecnología banda ancha mayorista (Neba Local, NEBA cobre...)
Unidades	Unidades
Líneas o Accesos (Núm)	Dato de Líneas o Accesos Unidades
Estaciones Base (Núm)	Dato de Estaciones Base Unidades
Tasa de Penetración (%)	Dato de Tasa de Penetración % - UNICAMENTE PARA PROVINCIAS

Estructura de datos

		Líneas de telefonía fija	Líneas de banda ancha fija minorista	Líneas telefonía móvil	Abonados de televisión de pago	Accesos Instalados	Estaciones Base	Líneas NEBA
Parámetros	Año	X	X	X	X	X	X	X
	Provincia	X	X	X	X	X	X	X
	CCAA	X	X	X	X	X	X	X
	Servicio	Telefonía fija	Banda ancha fija	Telefonía móvil	Servicios audiovisuales	Accesos instalados	Estaciones Base	Banda ancha fija mayorista
	Concepto	Líneas	Líneas	Líneas	Abonados	Nº de accesos por tipo de acceso	Nº de estaciones base en redes de acceso móvil	Nº de accesos por modalidad de acceso y equipamiento xDSL
	Operador	X	X	X				
	Tecnología de Acceso BAF		X					
	Tecnología de Acceso					X		
	Tipo de Estaciones Base						X	
	Tipo de BA Mayorista							X
Indicadores	Líneas o Accesos (Núm)	X	X	X	X	X		X
	Estaciones Base (Núm)						X	
	Tasa de Penetración (%)	X	X	X	X	X		X

En la tabla anterior se indica la estructura de la información y qué parámetros e indicadores se utilizan para cada variable representada en las estadísticas provinciales de Telecomunicaciones.

Por ejemplo, si quisiésemos extraer información de los “Abonados de televisión de pago” deberíamos seguir los siguientes pasos:

- Seleccionar el servicio de “Servicios audiovisuales” y el concepto de “Abonados”
- Disponemos información en los siguientes parámetros: Año, provincia y CCAA (ver tabla anterior).
- Seleccionar los indicadores de “Líneas o Accesos (Núm)” para el número total de líneas o “Tasa de Penetración (%)” para obtener la tasa de penetración.