

¿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

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, tipo de contrato, tecnología de acceso, mes, 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, ingresos, tráficos, penetraciones, portabilidades, clientes, 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 del de líneas.

Diccionario de datos

Columna	Descripción
mes	Indica el mes
País	País
Servicio	Indica el servicio
Concepto	líneas, portabilidades
Operador	Nombre del operador
Segmento	Parámetro de segmento (Residencial o negocios)
Tecnología de Acceso	Parámetro de tecnologías de acceso del servicio de "Banda ancha fija minorista" (ADSL; HFC; FTTH, ...)
Contrato	Parámetro de tipo de contrato (Pospago o prepago)
Terminal de Telefonía Móvil	Parámetro para el servicio de "telefonía móvil", (asociado a máquina o es un terminal)
Línea BA Móvil	Parámetro para el servicio de "Banda ancha móvil" (línea de voz y datos o línea exclusiva de datos).
Tipo de BA Mayorista	Parámetro el servicio de banda ancha fija mayorista (NEBA Local, Bucle compartido, etc...)
Tipo de Portabilidad	Tipo de portabilidad (importados o exportados)
unidades	Unidades en que se miden los indicadores
Líneas (Núm)	Dato de Líneas
Tasa de Penetración (%)	Dato de Tasa de Penetración
Portabilidades (Núm)	Dato de Portabilidades

Estructura de datos

		Líneas de telefonía fija	Portabilidades telefonía fija	Líneas telefonía móvil	Líneas máquina (M2M) telefonía móvil	Portabilidades móvil	Líneas Banda ancha fija minorista	Líneas banda ancha móvil	Líneas banda ancha fija mayorista
Parámetros	mes	X	X	X	X	X	X	X	X
	País	X	X	X	X	X	X	X	X
	Servicio	Telefonía fija	Telefonía fija	Telefonía móvil	Telefonía móvil	Telefonía móvil	Banda ancha fija minorista	Banda Ancha móvil	Banda ancha fija mayorista
	Concepto	Líneas	Portabilidades	Líneas	Líneas	Portabilidades	Líneas	Líneas	Líneas
	Operador	X		X	X	X	X	X	
	Segmento	X							
	Tecnología de Acceso						X		
	Contrato			X	X			X	
	Terminal de Telefonía Móvil			Sin tipo de terminal	Asociados a máquinas				
	Línea BA Móvil							X	
	Tipo de BA Mayorista								X
	Tipo de Portabilidad					X			
Indicadores	Líneas (Núm)	X		X	X		X	X	X
	Tasa de Penetración (%)	X		X	X		X		
	Portabilidades (Núm)		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 mensuales de Telecomunicaciones.

Por ejemplo, si quisiésemos extraer la información de las “Líneas de banda ancha minorista”, deberíamos seguir los siguientes pasos:

- Seleccionar el servicio de “Banda ancha fija minorista” y el concepto de “Líneas”
- Disponemos información en los siguientes parámetros: mes, operador y Tecnología de acceso.
- Seleccionar los indicadores de “Líneas (Núm)” para el número total de líneas o “Tasa de Penetración (%)” para obtener la tasa de penetración.