

Government Data Analysis (GOVEDA)

Unidad de científicos
de datos

Diciembre, 2017
dnp.gov.co



Análisis exploratorio de datos abiertos en Colombia

Objetivo General: Caracterizar la información disponible en los portales de datos abiertos

La librería desarrollada puede:

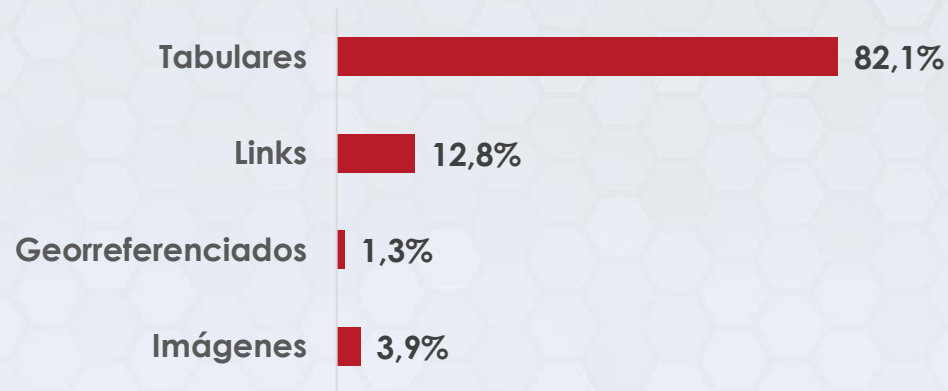
- Realizar análisis de portales CKAN y Socrata.
- Generar reporte básicos con la caracterización de los datos disponibles en los portales de datos abiertos
- Desplegar un reporte básico de los datos abiertos por tags, temáticas o entidades
- Exportar o extraer los datos en formatos para su consumo y análisis

Esto es posible por el desarrollo de scripts para análisis de portales de datos abiertos

Resultados para Colombia



Tipo de información



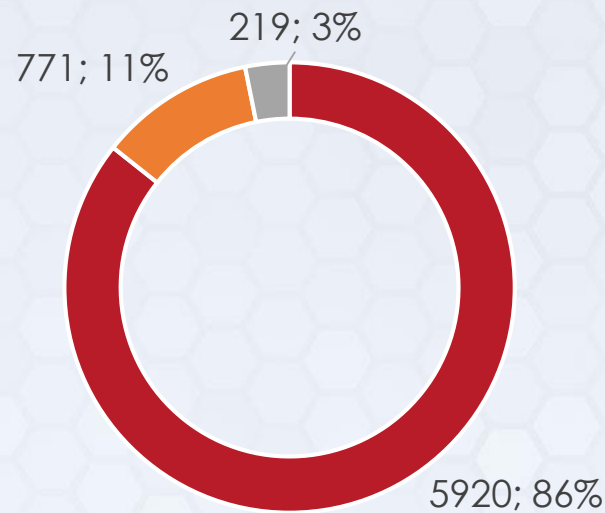
Número de Ficheros



Generalidades

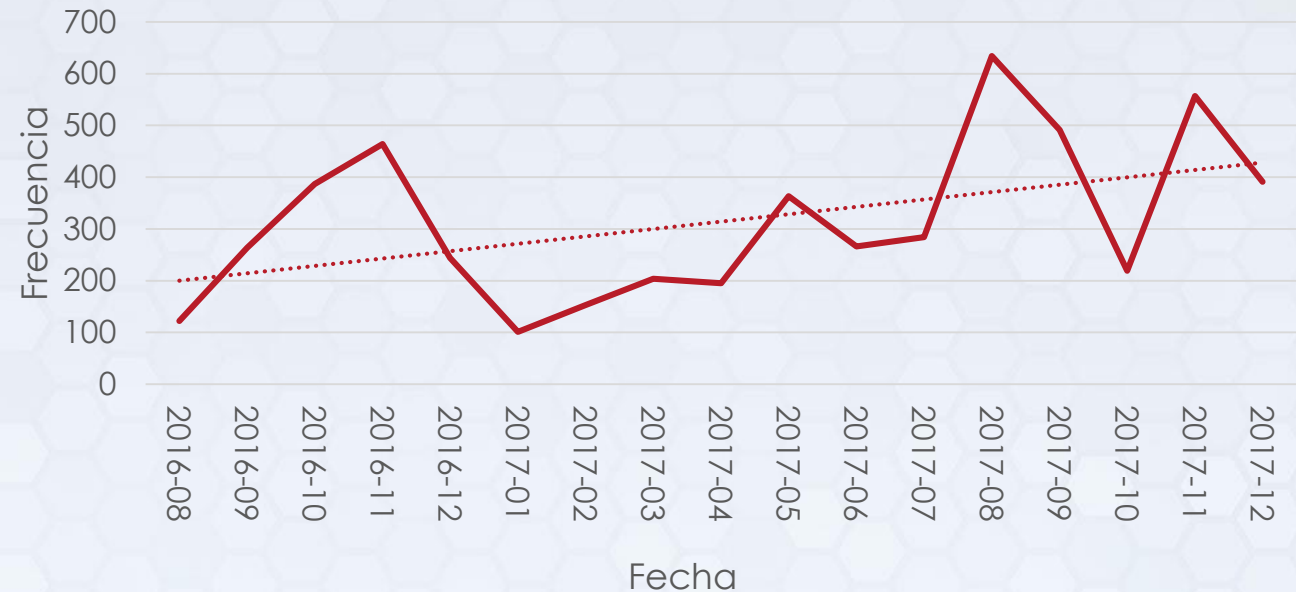
- ✓ El portal de datos abiertos aloja más de **6911 de ficheros de datos** (tablas, datos geográficos y links a información estratégica de la entidad)
- ✓ Información con el número de visitas, de descargas, la entidad que publica los datos, la temática entre otras características puede obtenerse vía una API (goveda),

6911 ficheros



■ tabular ■ href ■ geo

Ficheros Publicados



Se excluye los ficheros publicados en Junio y Julio de 2016 (4 y 1568 ficheros), ya que estos valores son atípicos,

Entidades

- ✓ **1635 entidades** han publicado a la fecha en el portal de datos abiertos,
- ✓ Hay 377 ficheros sin entidad identificada (5,5% de todos los 6911 ficheros publicados a la fecha),

Nro, de ficheros por entidad	Frec,	% entidades	Acum,
1	904	55,3	55,3
2	224	13,7	69,0
3	129	7,9	76,9
4	76	4,6	81,5
5	50	3,1	84,5
> 5	252	15,5	100

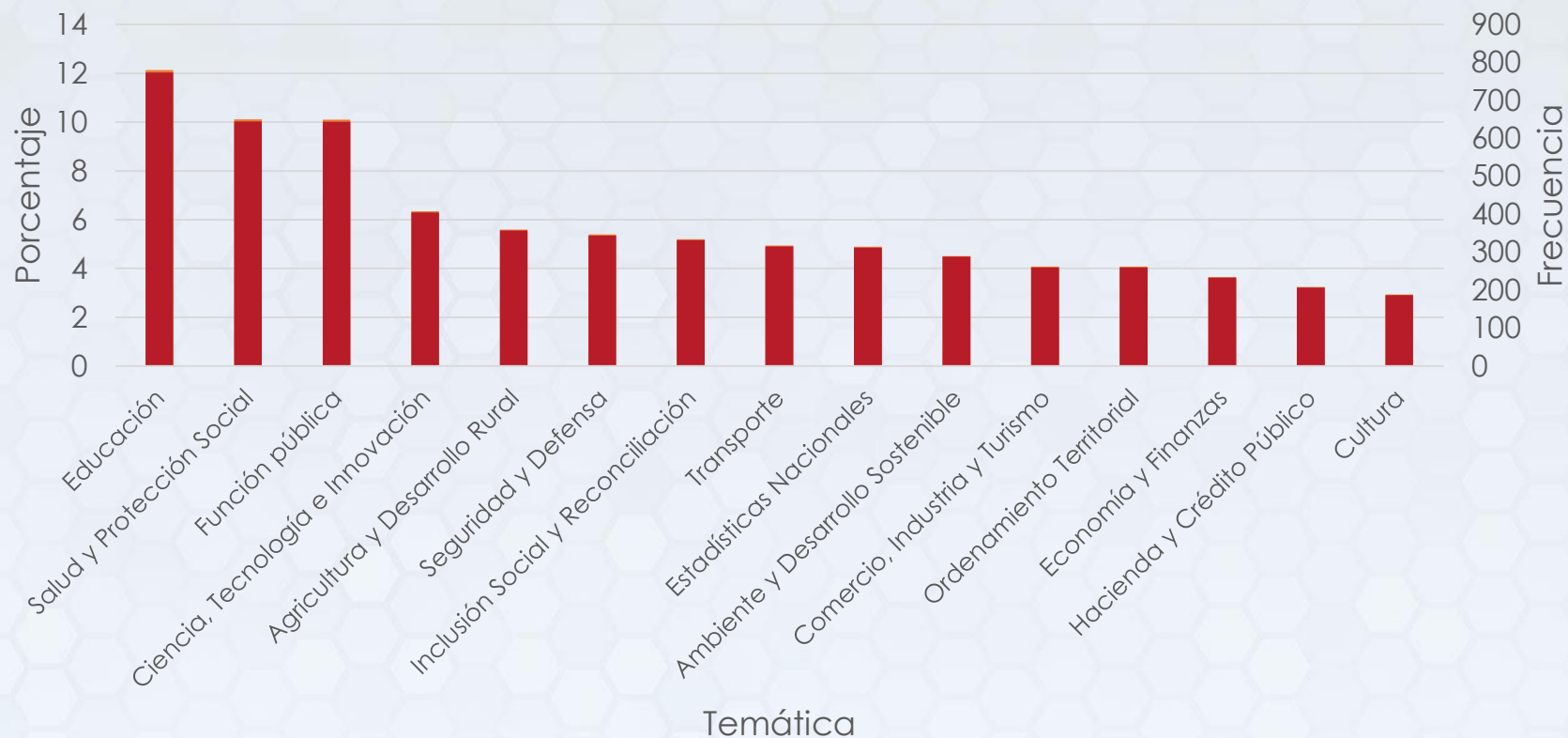
- ✓ El 55,3 % de las entidades publican sólo un fichero,
- ✓ El 84,5 % de las entidades publican 5 ficheros o menos,

Entidad	Nro, Ficheros	Porcentaje
MINTIC	261	4,0
Policía Nacional	229	3,5
DANE	160	2,4
Aguas y aguas de Pereira	131	2,0
Alcaldía de Cali	113	1,7
Alcaldía de Chía	110	1,7

Entidades con más de 100 ficheros publicados

Ficheros por temática

- ✓ El **87,2** de los ficheros tratan de las **15 siguientes temáticas** (de un total de 26 temáticas). Para ver en más detalle las otras temáticas ver: [Tabla.xlsx](#)
- ✓ Se destacan educación, salud y protección social, y Función Pública con cerca de la tercera parte de los ficheros.



Nota: Hay 459 ficheros sin temática identificadas, 74 ficheros de DataJam's. El análisis por temática se realiza sobre los 6378 ficheros restantes.

Número de visitas

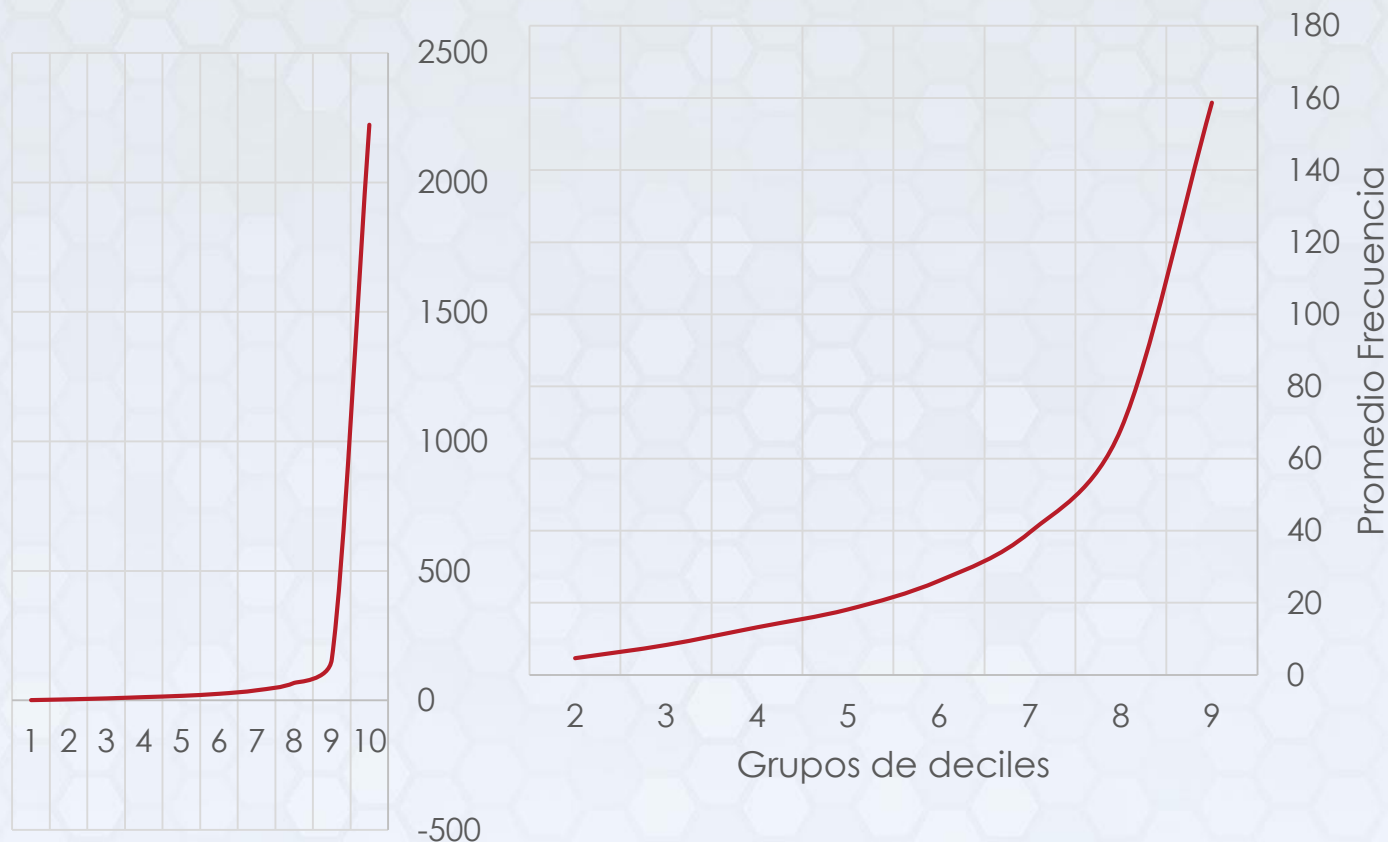
Existen cerca de un 10 % de ficheros con más de 1000 vistas, **la mediana del número de visitas es de 141** (el promedio es de 991,8). Los ficheros con más vistas son:



Entidades	Frec
INSTITUTO NACIONAL DE VIGILANCIA DE MEDICAMENTOS Y ALIMENTOS INVIMA	387223
GOBERNACIÓN DE BOYACÁ	40048
MINISTERIO DE SALUD Y PROTECCIÓN SOCIAL	36620
MINISTERIO DE TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES	35868
DEPARTAMENTO ADMINISTRATIVO DE CIENCIA, TECNOLOGÍA E INNOVACIÓN	32859
ALCALDÍA MUNICIPAL DE TULUÁ EN VALLE DEL CAUCA	28325
ALCALDÍA MUNICIPAL DE TUNJA EN BOYACÁ	25164
SUPERINTENDENCIA DE LA ECONOMÍA SOLIDARIA	23988
DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA	20857
MINISTERIO DE EDUCACIÓN NACIONAL	20297

Número de descargas

Existen cerca de un 10 % de ficheros con más descargas, **la mediana del número de descargas es de 21** (el promedio es de 254,4,8). Los ficheros con más vistas son:



Entidades	Frec
SUPERINTENDENCIA DE LA ECONOMÍA SOLIDARIA	54410
INSTITUTO NACIONAL DE VIGILANCIA DE MEDICAMENTOS Y ALIMENTOS INVIMA	21130
POLICÍA NACIONAL	16234
REGISTRADURÍA NACIONAL DEL ESTADO CIVIL	13507
CONTRALORIA GENERAL DE ANTIOQUIA	13170
MINISTERIO DE EDUCACIÓN NACIONAL	11764
UNIDAD ADMINISTRATIVA ESPECIAL DE CATASTRO DISTRITAL	11462
MINISTERIO DE TECNOLOGÍAS DE LA INFORMACIÓN Y LAS COMUNICACIONES	10497
ALCALDÍA MUNICIPAL DE BUCARAMANGA EN SANTANDER	8597
AGENCIA NACIONAL DE CONTRATACIÓN PÚBLICA - COLOMBIA COMPRA EFICIENTE	8446