

OPEN DATA

GFT

Esteban Chiner

Curso 2021/2022

Fecha 04/11/2021

Esteban Chiner



- **Computer Engineer** from the Universitat de València & **Master on Software Engineering** from Universitat Politècnica de València
- Working in GFT as a **Senior Architect** specialized in Big Data and Blockchain
- Part of **Offering Development** team, leading the “**Data**” and “**DLT & Crypto**” offerings
- **Lecturer**
 - Big Data & Analytics Master (**UPV**) – “Big Data Architectures in Financial Services”
 - Data Analytics for Enterprises (**EDEM**) + several seminars on Blockchain
 - Seminaris d'Empresa (**UPC**) – “Developing decentralized apps with Ethereum”
- **Hadoop Certified** in developer, administration and HBase
- Paper published: “**A Big Data Financial Information Management Architecture for Global Banking**”



[@echiner](https://twitter.com/echiner)



[echiner](https://www.linkedin.com/company/echiner)



esteban.chiner@gft.com

Agenda

1. Introduction

- Introduction and types

2. Technical Background

- Data formats
- Delivery mechanisms
- Frequency

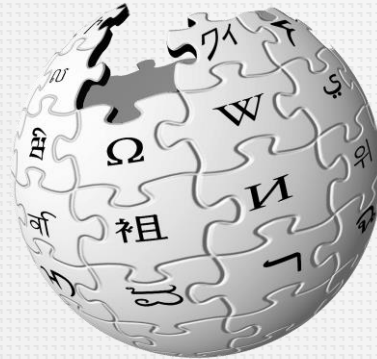
3. Datasources

- Government
- Statistics
- Public Datasets
- Commercial

Introduction

Introduction

“**Open Data** is the idea that some **data** should be **freely available to everyone** to use and republish as they wish, **without restrictions** from copyright, patents or other mechanisms of control.”



Types of Open Data



Provided by different governments levels (local, regional, national or supra national).

Pure “Open Data”.

Examples: Ajuntament de Valencia, European Data Portal, Data.gov...

Government



Can be considered a sub-type of government open data.

Provides economic or statistical information.

Examples: INE, World Bank Open Data...

Statistics



Usually used for testing purposes or in Artificial Intelligence training.

Cannot be really considered a data source, but a single dataset.

Examples: Google Dataset Search (Kaggle, DataCite...)

Public Datasets

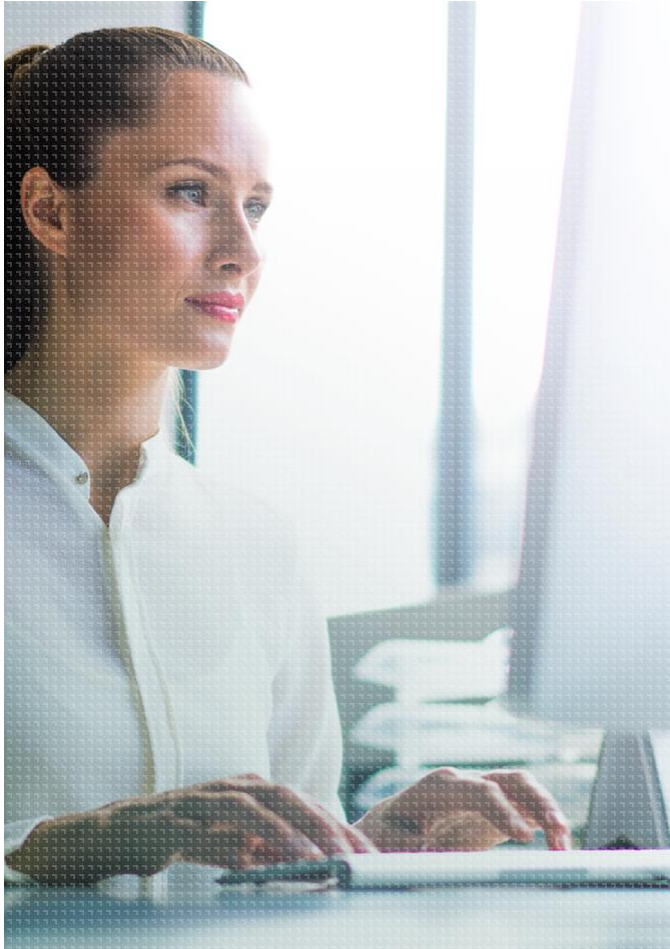


Data is provided for commercial purposes.

Payment per use (calls, datasets, etc.), although in many cases they provide a free tier.

Examples: Social Networks (Twitter, Reddit...), Financial Data (IEX, Morning Star...)

Commercial



What other examples do you know?

Think about other examples of Open Data in the described areas:

- Government
- Statistics
- Public datasets
- Commercial

Technical Background

Data formats (I)

CSV (Comma Separated Values)

Header (optional)

"LatD",	"LatM",	"LatS",	"NS",	"LonD",	"LonM",	"LonS",	"EW",	"City",	"State"
41,	5,	59,	"N",	80,	39,	0,	"W",	"Youngstown",	OH
42,	52,	48,	"N",	97,	23,	23,	"W",	"Yankton",	SD
46,	35,	59,	"N",	120,	30,	36,	"W",	"Yakima",	WA
42,	16,	12,	"N",	71,	48,	0,	"W",	"Worcester",	MA
43,	37,	48,	"N",	89,	46,	11,	"W",	"Wisconsin Dells",	WI

Content

Data formats (II)

JSON (JavaScript Object Notation)

```
{
  "glossary": {
    "title": "example glossary",
    "GlossDiv": {
      "title": "S",
      "GlossList": {
        "GlossEntry": {
          "ID": "SGML",
          "SortAs": "SGML",
          "GlossTerm": "Standard Generalized Markup Language",
          "Acronym": "SGML",
          "Abbrev": "ISO 8879:1986",
          "GlossDef": {
            "para": "A meta-markup language, used to create markup languages such as DocBook.",
            "GlossSeeAlso": ["GML", "XML"]
          },
          "GlossSee": "markup"
        }
      }
    }
  }
}
```

Column Name

Column Value (String)

Column Value (Array)

Data formats (III)

XML (Extensible Markup Language)

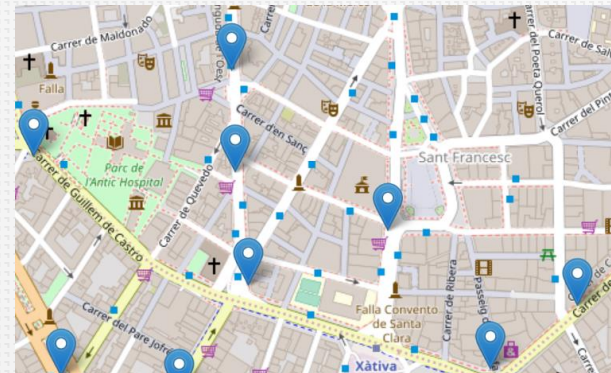
```
<!DOCTYPE glossary PUBLIC "-//OASIS//DTD DocBook V3.1//EN">
<glossary><title>example glossary</title>
  <GlossDiv><title>S</title>
    <GlossList>
      <GlossEntry ID="SGML" SortAs="SGML">
        <GlossTerm>Standard Generalized Markup Language</GlossTerm>
        <Acronym>SGML</Acronym>
        <Abbrev>ISO 8879:1986</Abbrev>
        <GlossDef>
          <para>A meta-markup language, used to create markup languages such as DocBook.</para>
          <GlossSeeAlso OtherTerm="GML">
            <GlossSeeAlso OtherTerm="XML">
              </GlossDef>
            <GlossSee OtherTerm="markup">
              </GlossEntry>
            </GlossList>
          </GlossDiv>
        </glossary>
```

Data formats (III)

Geographical Information Formats

- **WMS** (Web Map Service)
 - International standard by OGC
 - Serves georeferenced map images
- **GML** (Geography Markup Language)
 - Developed by OpenGIS
 - Based on XML
- **KML** (Keyhole Markup Language)
 - Used for Google Earth
 - Similar to GML
 - Can be compressed as KMZ

```
<Placemark>
  <name>Alameda, 51/ Menorca</name>
  <styleUrl>#style_symbol_CIRCULO_100.0_ff9b93ff</styleUrl>
  <ExtendedData>
    <Data name="emplazamie">
      <displayName>emplazamie</displayName>
      <value><![CDATA[Alameda, 51/ Menorca]]></value>
    </Data>
    <Data name="modelo">
      <displayName>modelo</displayName>
      <value><![CDATA[ALBUFERA]]></value>
    </Data>
  </ExtendedData>
  <Point>
    <coordinates>
      -0.34840107,39.45662252
    </coordinates>
  </Point>
</Placemark>
```



Data formats (IV)

News feeds (Atom & RSS)

- Described in XML format
- Delivered via HTTP
- Many readers available (incl. Browser)
- Basically same features, although Atom more advanced

```
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0">

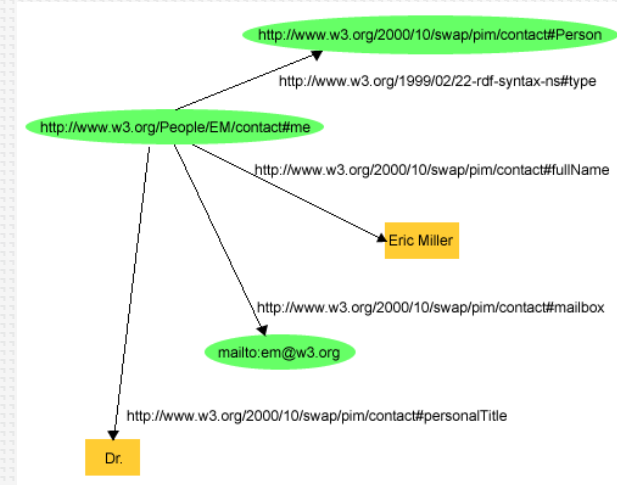
<channel>
  <title>W3Schools Home Page</title>
  <link>https://www.w3schools.com</link>
  <description>Free web building tutorials</description>
  <item>
    <title>RSS Tutorial</title>
    <link>https://www.w3schools.com/xml/xml_rss.asp</link>
    <description>New RSS tutorial on W3Schools</description>
  </item>
  <item>
    <title>XML Tutorial</title>
    <link>https://www.w3schools.com/xml</link>
    <description>New XML tutorial on W3Schools</description>
  </item>
</channel>

</rss>
```


Data formats (V)

RDF (Resource Description Framework)

- Defined by the W3C, used for modelling web resources
- Similar to Entity-Relationship or Class Diagrams
- Uses **SPARQL** as Query Language
- Can be serialized in different formats (Turtle, JSON, XML...)



```
@prefix eric:    <http://www.w3.org/People/EM/contact#> .
@prefix contact: <http://www.w3.org/2000/10/swap/pim/contact#> .
@prefix rdf:     <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .

eric:me contact:fullName "Eric Miller" .
eric:me contact:mailbox <mailto:e.miller123(at)example> .
eric:me contact:personalTitle "Dr." .
eric:me rdf:type contact:Person .
```

Delivery mechanisms



DIRECT DOWNLOAD

A

This is the most common option for Open Data and public datasets, since they just provide downloadable content.



REST API

C

Representational State Transfer (REST) defines how web services should be defined. An URI (endpoint) represents a service and can come in any format (usually JSON).



REAL-TIME

B

Data is provided in real-time as an unbounded dataset (stream of data). It can be done via REST endpoint, messaging (e.g. broker, websocket...) or callback (e.g. webhook).



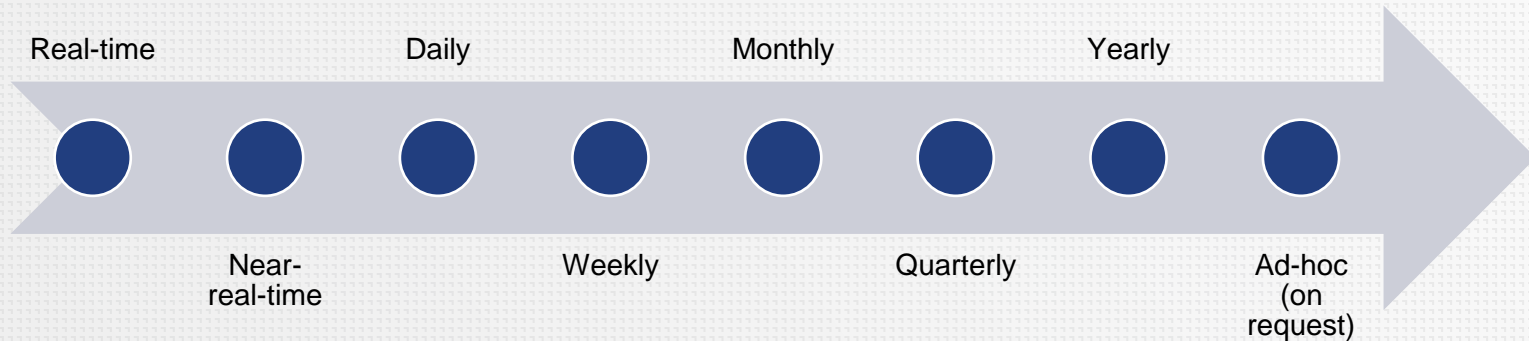
FEEDS

D

HTTP call which provide a result which gets updated more or less frequently. Examples being RSS and Atom.

Frequency

- Depending on the **nature of the data** and the **frequency it is updated**, it can be delivered in many different time spans.



Data Sources

Government

Government – Ajuntament de València



Ajuntament de València AJUNTAMENT LA CIUTAT TRÀMITS ATENCIÓ CIUTADANA ACTUALITAT

Què estàs buscant? 🔍

Ajuntament Dades Obertes

- DADES OBERTES
- CATÀLEG DE DADES
- QUÈ SON?
- QUADRE D'INDICADORS
- APIS DISPONIBLES
- APLICACIONS
- COL·LABORA

🏠 > AJUNTAMENT > GOVERN OBERT > DADES OBERTES

Dades Obertes

Qualsevol ciutadà o empresa pot analitzar, reutilitzar i redistribuir aquestes dades generant nous serveis. Això també permet que l'administració pública millore en transparència i govern obert, al mateix temps que fomenta la generació de riquesa mitjançant una gestió intel·ligent dels recursos (govern intel·ligent).

-  Societat i Benestar
-  Medi Ambient
-  Transport
-  Urbanisme i Infraestructures

<https://www.valencia.es/dadesobertes/va/>

Government – Ajuntament de València

🏠 > AJUNTAMENT > GOVERN OBERT > DADES OBERTES > CATÀLEG DE DADES



Medi Ambient



Societat i Benestar



Transport



Urbanisme i
Infraestructures



Salut



Turisme



Sector Públic



Cultura i Oci



Comerç



Economia



Hisenda



Ciència i
tecnologia



Educació



Seguretat



Vivenda

Buscar en el Catàleg de Dades...



<https://www.valencia.es/dadesobertes/va/>

Government – Ajuntament de València

Aparcament Mobilitat Reduïda

Aparcaments per a vehicles de persones amb mobilitat reduïda. Plazas: Nombre de places d'aparcament

Formats de distribució: WFS, GEOJSON, SHP, GML, WMS, KML, KMZ, CSV

Ajuntament de València

Visites 58

Societat i Benestar Transport

Grau d'obertura
★★★★☆

Estacions de Valenbisi

Localització de les estacions de valenbisi. Name: Nom de l'estació de valenbisi Number: Número de l'estació. Address: direcció. Open: Si està operativa o no. Available: Bicicletes...

Formats de distribució: WFS, GEOJSON, SHP, GML, WMS, KML, KMZ, CSV

Ajuntament de València

Visites 138

Transport Turisme

Grau d'obertura
★★★★☆

Carril Bici

Itineraris ciclistes de la ciutat de València

Formats de distribució: WFS, GEOJSON, SHP, GML, WMS, KML, KMZ, CSV

Ajuntament de València

Visites 160

Transport Turisme

Grau d'obertura
★★★★☆

Bol·lards extraïbles

Bol·lards extraïbles amb clau estàndard aprovada pel Servei de Bombers

Formats de distribució: GML, GEOJSON, KML, KMZ, SHP, WFS, WMS, CSV

Ajuntament de València

Visites 31

Transport

Grau d'obertura
★★★★☆

ETIQUETES

circulació (4)
parking (4)
aparcament (3)
vehicle (2)
vehículo (2)

+ Expandir etiquetes

FREQUÈNCIA D'ACTUALITZACIÓ

Diària (25)
Setmanal (36)
Mensual (15)
Trimestral (10)
Semestral (4)

+ Expandir

DARRERA ACTUALITZACIÓ

1/1/2013 31/10/2021

FORMATS

CSV (120)
SHP (100)
WMS (100)
WFS (99)
GEOJSON (98)

+ Expandir formats

GRAU D'OBTURA

★★★★☆
★★★★☆
★★★★☆
★★★★☆
★★★★☆

<https://www.valencia.es/dadesobertes/va/>

Government – Ajuntament de València

Carril Bici

Itineraris ciclistes de la ciutat de València

161 VISITES



METADADES

Sectors	Transport, Turisme
Condicions d'ús	Atribución 4.0 Internacional (CC BY 4.0)
Grau d'obertura	★★★★☆
Data de creació	10/12/2019 13:46:58
Freqüència d'actualització	Trimestral
Organisme	



**AJUNTAMENT
DE VALÈNCIA**

DISTRIBUCIONS

	Carril Bici Itineraris ciclistes de la ciutat de València	28 
	Carril Bici Itineraris ciclistes de la ciutat de València	20 
	Carril Bici Itineraris ciclistes de la ciutat de València	58 
	Carril Bici Itineraris ciclistes de la ciutat de València	7 

<https://www.valencia.es/dadesobertes/va/>

Government – Spanish Government

INICIATIVA APORTA ▾ CATÁLOGO DE DATOS ▾ IMPACTO ▾ SECTORES ▾ ACTUALIDAD ▾ INNOVACIÓN ▾ INTERACTÚA ▾

CONJUNTOS DE DATOS **API** PUNTO SPARQL

Catálogo de datos.

 CIENCIA Y TECNOLOGÍA (1869)	 COMERCIO (1064)	 CULTURA Y OCIO (2251)	 DEMOGRAFÍA (8053)	 DEPORTE (364)	 ECONOMÍA (4317)	 EDUCACIÓN (3145)	 EMPLEO (4033)	 ENERGÍA (555)
 HACIENDA (2448)	 INDUSTRIA (579)	 LEGISLACIÓN Y JUSTICIA (4573)	 MEDIO AMBIENTE (6007)	 MEDIO RURAL (1226)	 SALUD (4180)	 SECTOR PÚBLICO (10327)	 SEGURIDAD (756)	 SOCIEDAD Y BIENESTAR (7622)
<div>     </div> <div> TRANSPORTE (1290) TURISMO (2774) URBANISMO E INFRAESTRUCTURAS (1751) VIVIENDA (735) </div>								

<https://datos.gob.es>

Government – Spanish Government

dataset

GET	/catalog/dataset	Finds all datasets
GET	/catalog/dataset/{id}	Finds a dataset by the URI Identifier
GET	/catalog/dataset/title/{title}	Finds datasets by title
GET	/catalog/dataset/publisher/{id}	Finds datasets by publisher
GET	/catalog/dataset/theme/{id}	Finds datasets by theme
GET	/catalog/dataset/format/{format}	Finds all datasets that have a distribution with a format that you use as parameter
GET	/catalog/dataset/keyword/{keyword}	Finds all datasets that have a keyword that you use as parameter

<https://datos.gob.es/es/apidata#!/dataset/findDatasetById>

<https://datos.gob.es>

Government – European Union

data.europa.eu

Español (es)

Buscar contenido del si



Datos ^

Estudios v

data.europa academy

Noticias v

Contacto

Conjuntos de datos

Búsqueda SPARQL

Estadísticas

Calidad de los metadatos

URI persistentes

Vocabularios de la UE

Buscar conjuntos de datos



Agricultura, pesca,
silvicultura y
alimentos



Asuntos
internacionales



Ciencia y tecnología



Economía y finanzas



Educación, cultura y
deportes



Energía



Gobierno y sector
público



Justicia,
ordenamiento jurídico
y seguridad pública



Medio ambiente



Población y sociedad



Regiones y ciudades



Salud



Transporte

Buscar conjuntos de datos



<https://data.europa.eu/es>

Government – European Union

Prefijos



```
1 SELECT (count(*) AS ?count)
2 WHERE { { ?s a dcat:Dataset } }
3 LIMIT 100
4 |
```

Vista previa de los datos de resultados

Formato

Limite

```
1 <rdf:RDF
2   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3   xmlns:rs="http://www.w3.org/2001/sw/DataAccess/tests/result-set#"
4   xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
5 <rs:ResultSet>
6   <rs:size rdf:datatype="http://www.w3.org/2001/XMLSchema#int"
7     >1</rs:size>
8   <rs:solution rdf:parseType="Resource">
9     <rs:binding rdf:parseType="Resource">
10      <rs:value rdf:datatype="http://www.w3.org/2001/XMLSchema#integer"
11        >917309</rs:value>
12      <rs:variable>count</rs:variable>
13    </rs:binding>
14  </rs:solution>
```

Descargar

<https://data.europa.eu/es>

Government – European Union

Query ✕ +

⚙️ https://data.europa.eu/sparql

```

1 PREFIX dcat: <http://www.w3.org/ns/dcat#>
2 PREFIX odp: <http://data.europa.eu/euodp/ontologies/ec-odp#>
3 PREFIX dct: <http://purl.org/dc/terms/>
4 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
5 PREFIX foaf: <http://xmlns.com/foaf/0.1/>
6
7 SELECT * WHERE { ?d a dcat:Dataset } LIMIT 10
8

```

🔗

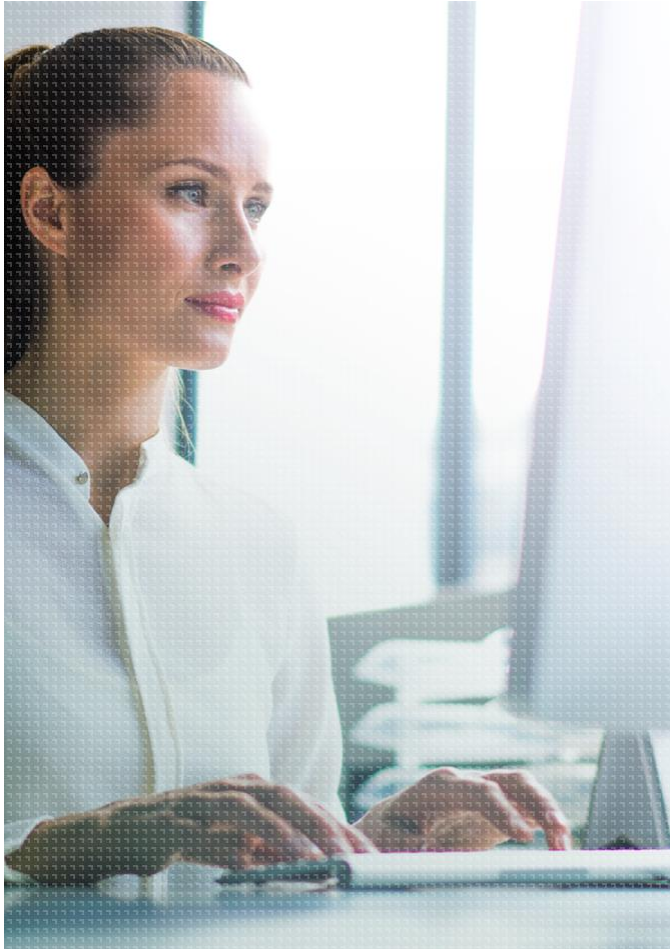
Table
Response
10 results in 0.176 seconds

```

1 {
2   "head": {
3     "link": [],
4     "vars": [
5       "d"
6     ]
7   },
8   "results": {
9     "distinct": false,
10    "ordered": true,
11    "bindings": [
12      {
13        "d": {
14          "type": "uri",
15          "value": "http://data.europa.eu/88u/dataset/5ig8huzrleocpdue37xow"
16        }
17      },
18      {
19        "d": {
20          "type": "uri",
21          "value": "http://data.europa.eu/88u/dataset/honey-bee-seasonal-mortality-2012-2014-epilobee-analysis"
22        }
23      }
24    ]
25  }
26 }

```

<https://data.europa.eu/es>



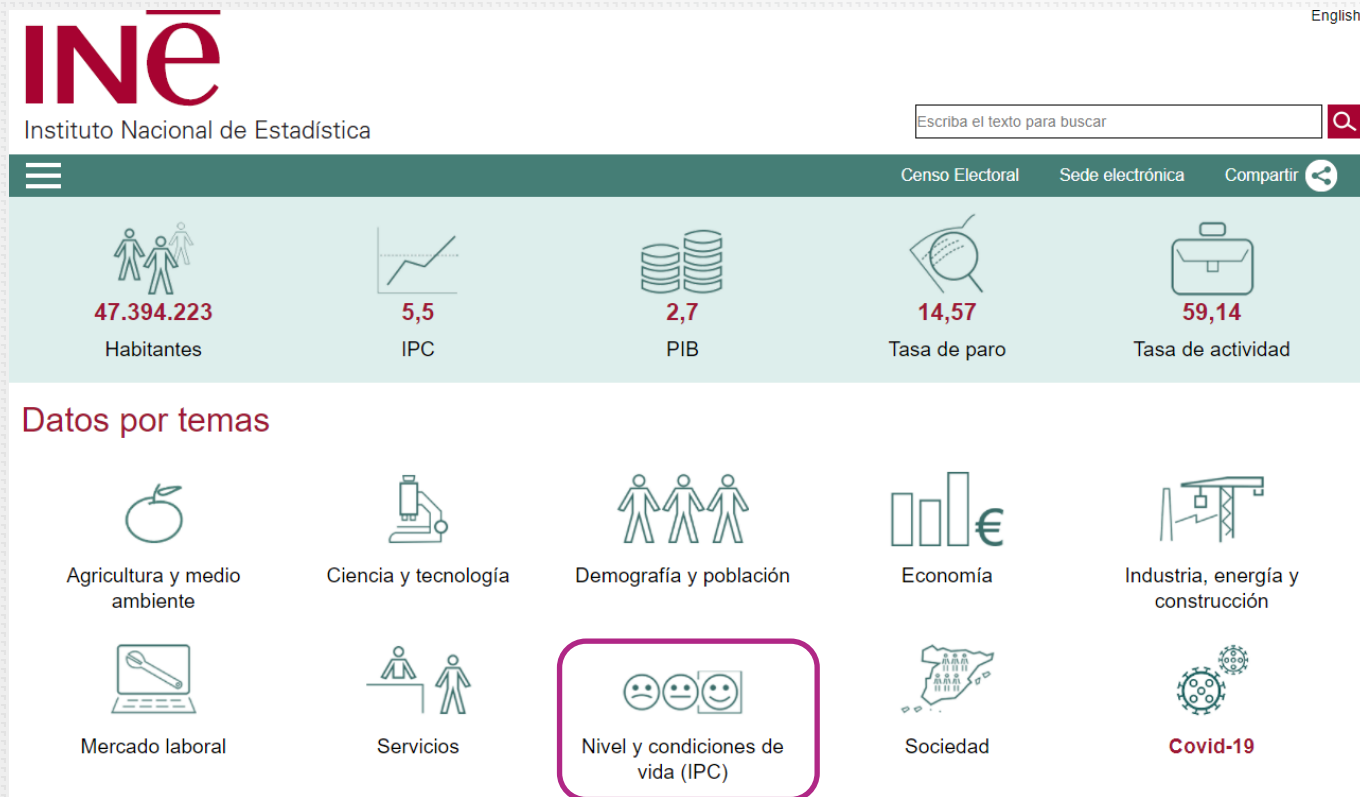
Analysing a dataset

- Search for **a couple of datasets** from two different sources of government Open Data
- Download in **CSV** format
- Load into **Excel** and “**analyse**”
- Share with the class

Data Sources

Statistics

Statistics – INE (Instituto Nacional de Estadística)



<https://www.ine.es/>

Statistics – INE (Instituto Nacional de Estadística)

INEbase / Nivel y condiciones de vida (IPC)

Estadísticas A...Z

Lista completa de
operaciones

Agricultura y medio
ambiente

Ciencia y tecnología

Demografía y población

Economía

Industria, energía y
construcción

Mercado laboral

Servicios

Nivel y condiciones de
vida (IPC)

Sociedad

Covid-19

Nivel y condiciones de vida (IPC)

Índices de precios de consumo y vivienda

Operaciones estadísticas que el INE elabora de forma periódica

Últimos datos

Índice de precios de consumo

Avance. Octubre 2021

Índice de precios de consumo armonizado

Avance. Octubre 2021

Índice de precios de vivienda

Trimestre 2/2021

Operaciones elaboradas por otros organismos del sistema estadístico nacional

Estadística de Valor Tasado de Vivienda

Estadística de Precios del Suelo

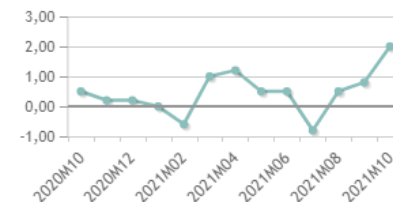
Estadística de Precios Públicos Universitarios

Condiciones de vida

Índice de precios de consumo. Base 2016 - Avance. Octubre 2021

	Variación mensual	Variación anual
Indice general	2,0	5,5
Bienes COVID-19	1,1	1,3
Servicios COVID-19	3,6	16,2
Alimentos y bebidas no alcohólicas	1,3	1,6
Alimentos sin elaboración	2,5	0,8
Alimentos elaborados	0,5	1,8

Indice general. Variación mensual



<https://www.ine.es/>

Statistics – Eurostat

Sign In | Register

Legal notice | RSS | Cookies | Links | Contact | English

Type a keyword, a publication title, a dataset title...

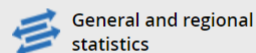
News | Data | Publications | About Eurostat | Help

European Commission > Eurostat

> Database

> Browse statistics by theme

STATISTICS BY THEME



General and regional
statistics



Economy and finance



Population and social
conditions



Industry, trade and
services



Agriculture and fisheries



Transport



Environment and
energy

Statistics Explained

Statistics A to Z

Government revenue, expenditure and main aggregates

Last update: 21-10-2021

Table Customization [about](#)

TIME		GEO									
+ Sector		+ National accounts indicator (ESA 2010)									
General government		Total general government expenditure									
GEO		2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
European Union - 27 countries		5,565,568.5	5,660,236.7	5,708,987.2	5,769,961.1	5,874,538.1	5,942,485.0	6,105,908.5	6,295,747.9	6,520,492.5	7,117,763.7
Euro area - 19 countries (from 2011)		4,842,911.9	4,918,492.3	4,960,345.4	5,012,823.5	5,095,520.6	5,163,760.7	5,289,225.4	5,436,975.4	5,622,014.7	6,129,127.1
Belgium		207,926.9	218,102.1	220,469.6	224,069.4	223,850.7	228,451.3	231,560.7	240,313.2	247,898.4	270,499.0 (P)
Bulgaria		27,338.4	28,312.2	31,057.0	36,329.2	38,172.9	33,155.3	35,724.8	40,622.8	42,737.3	45,139.0
Czechia		1,754,422.0	1,826,725.0	1,766,916.0	1,852,355.0	1,939,612.0	1,906,806.0	1,992,172.0	2,196,478.0	2,377,606.0	2,669,119.0
Denmark		1,042,167.0	1,098,247.0	1,077,153.0	1,093,954.0	1,110,377.0	1,106,149.0	1,108,531.0	1,139,013.0	1,147,822.0	1,244,789.0
Germany (until 1990 former GDR)		1,218,524.0	1,233,138.0	1,263,544.0	1,296,940.0	1,335,789.0	1,390,374.0	1,443,273.0	1,492,813.0 (P)	1,562,698.0 (P)	1,712,131.0 (P)
Estonia		6,241.4	7,031.5	7,266.3	7,571.1	8,154.5	8,568.1	9,351.4	10,184.3	10,936.0	12,313.1
Ireland		81,157.5	74,615.3	73,241.6	76,355.6	75,791.2	75,919.6	77,802.2	82,596.6	86,285.6	102,033.6
Greece		112,064.0	106,844.0	112,926.0	89,913.0	95,336.0	87,154.0	85,871.0	87,137.0	87,758.0	98,871.0
Spain		490,976.0	501,688.0	467,649.0	465,650.0	472,962.0	472,733.0	478,804.0	501,774.0	524,037.0	588,279.0

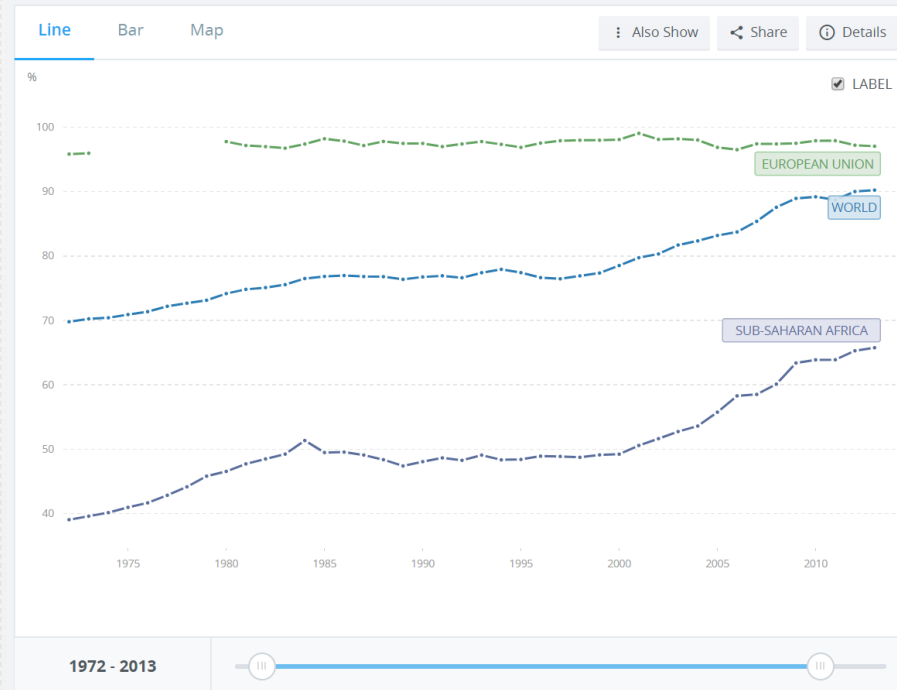
Statistics – World Bank



THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

AREAS:

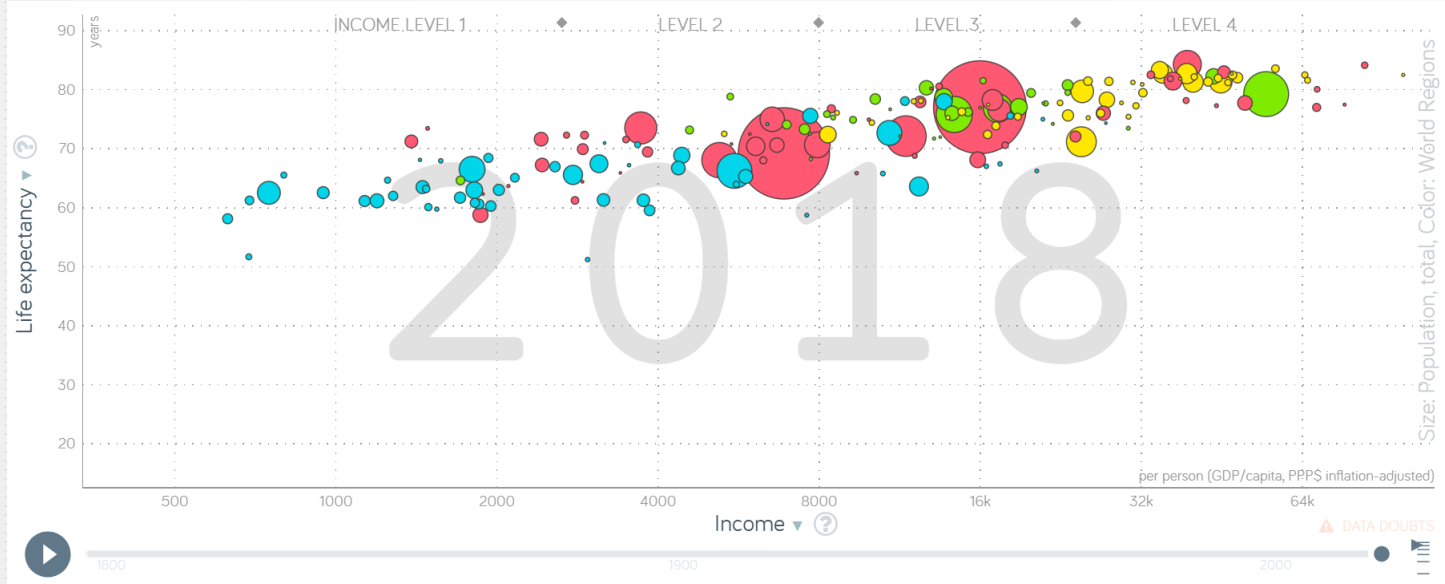
- Agriculture & Rural Development
- Aid Effectiveness
- Climate Change
- Economy & Growth
- Education
- Energy & Mining
- Environment
- External Debt
- Financial Sector
- Gender
- Health
- Infrastructure
- Poverty
- Private Sector
- Public Sector
- Science & Technology
- Social Development
- Social Protection & Labor
- Trade
- Urban Development



<https://data.worldbank.org/>

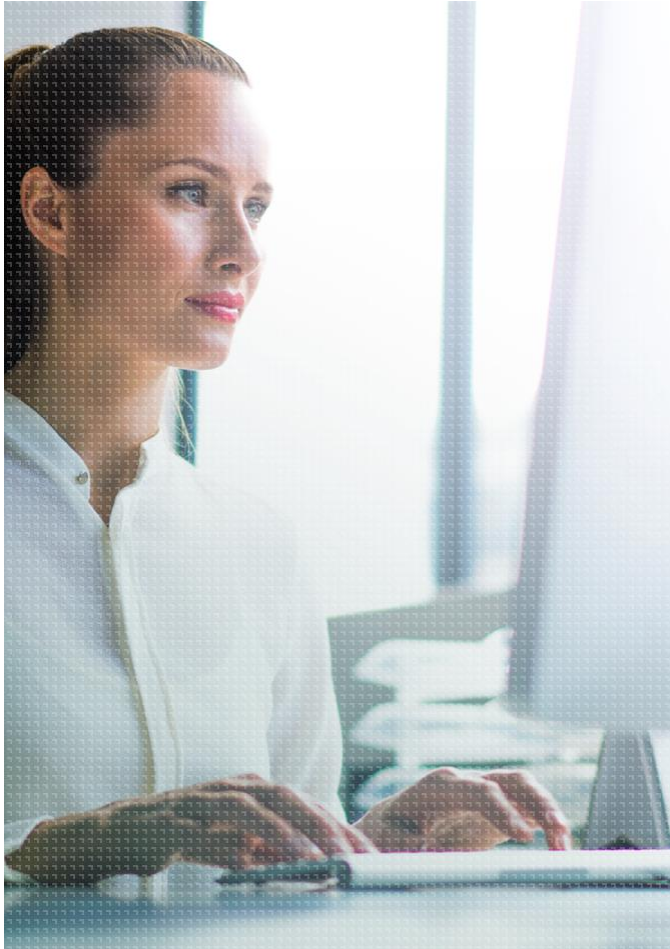
Statistics – Gapminder

- Great tool for **data crunching** and **visualization**



[https://www.gapminder.org/tools/#\\$chart-type=bubbles](https://www.gapminder.org/tools/#$chart-type=bubbles)

<https://www.gapminder.org/tools>



Socio-economic analysis

- In the **World Bank Data** site:
 - Compare the GDP Growth of “Low Income” vs “High Income” countries
 - Check the evolution of extreme poverty (1.90\$/day) over time
 - What is the fertility rate in India? Compare to China, Spain and the rest of the world
 - CO2 emissions → compare China, USA and European Union
 - Now change to CO2 emissions per capita
- In **Gapminder**:
 - Child Mortality vs Income → What is the evolution over time
 - Sustainability Development Index vs Income
 - Human Development Index vs Income
- (Optional) In **INE**, search for the evolution of the gap in GNP (PIB) between regions in Spain.
- Try to explain the results

Data Sources

Public Datasets

Public Datasources – Google Search

Google Dataset Search Beta

financial



Quarterly **Financial** Report: U.S. Corporations: Retained Earnings: All Information Industry
Net Lending (+) / Net Borrowing (-) (balance from **Financial** Account) as Direct Investment for Angola
United Arab Emirates BoP: Capital and **Financial** Account (CF)
Venezuela, RB - Global **Financial** Inclusion (Global Findex) Database 2017
Bangladesh BD: BOP: **Financial** Account: Foreign Direct Investment: Net Inflows
Ivory Coast CI: BoP: Capital Account: Acquisitions/Disposal of Non Produced Non **Financial** Assets
Jet Airways key **financial** figures 2013-2018
Net **financial** wealth by **financial** wealth and Total wealth band, all individuals, Great Britain July 2014 to July 2016
Bhutan BT: BOP: **Financial** Account: Foreign Direct Investment: Net Inflows
PVN-Annual report on AUDITED CONSOLIDATED **FINANCIAL** STATEMENTS

<https://toolbox.google.com/datasetsearch>

Public Datasources – Curated lists

70 free and amazing data sources for data visualization

- <https://bigdata-madesimple.com/70-amazing-and-free-data-sources-for-data-visualization/>

Big Data And AI: 30 Amazing (And Free) Public Data Sources For 2018

- <https://www.forbes.com/sites/bernardmarr/2018/02/26/big-data-and-ai-30-amazing-and-free-public-data-sources-for-2018/#1b8fed5f5f8a>

70 Amazing Free Data Sources You Should Know

- <https://www.kdnuggets.com/2017/12/big-data-free-sources.html>

Data Sources

Commercial

Commercial – Introduction

- There are many companies which provide data in **different models** (DaaS, pay per use, downloadable content, indirect payment...) and they are provided by **data vendors**.



Social Networks

Examples: Twitter, Reddit or Meetup



Economy & Financial

Examples: IEX Cloud or Quandl



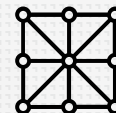
Simulated Data

Examples: PubNub on sensor data



Cloud Datasets

Examples: GCP Big Query datasets



Many Others

Examples: Crypto exchanges, Google Analytics

Commercial – IEX Cloud

- IEX Cloud is a **source of curated financial data**, providing institutional grade data, including fundamentals, ownership, international equities, mutual funds, options, **real-time data**, and alternative data – all in one **API**.

SCALE	GROW	LAUNCH
<p>\$699/me</p> <h2>499/mo</h2> <p>Billed annually Great for demanding applications and established businesses</p> <p>Select Scale</p> <ul style="list-style-type: none"> ✓ \$1 per 3,000,000 Pay as you go messages Start with ✓ 2,000,000,000 Messages per month ✓ Personal & Commercial use ✓ Enterprise Data ✓ 99.95% SLA ✓ Premium Support 	<p>\$59/me</p> <h2>49/mo</h2> <p>Billed annually Best for teams and developers that support users</p> <p>Select Grow</p> <ul style="list-style-type: none"> ✓ \$1 per 2,000,000 Pay as you go messages Start with ✓ 100,000,000 Messages per month ✓ Personal & Commercial use 	<p>\$19/me</p> <h2>9/mo</h2> <p>Billed annually Ideal for individual developers and students</p> <p>Select Launch</p> <ul style="list-style-type: none"> ✓ \$1 per 1,000,000 Pay as you go messages Start with ✓ 5,000,000 Messages per month ✓ Personal & Commercial use

<https://iexcloud.io>

Commercial – Twitter

- **Twitter Developer API** allows you to **publish and analyze** Tweets, **optimize ads**, and create unique **customer experiences**.
- Provides the following **functionality** via API:

Account & Users:

- List followers and Friends
- Search users
- Manage account
- Mute, block and report

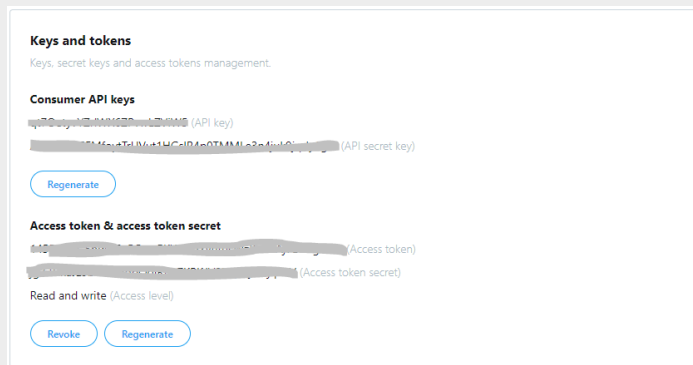
Tweets:

- Get user timeline
- Post tweets
- **Real-time tweets (sample)**
- Search

<https://developer.twitter.com>

Setting up Dev Twitter Account

- Sign up to Twitter Dev:
 - <https://developer.twitter.com>
- Go to “Apps”
- “Create an app” and once created, under “Keys and tokens” you will see the following:



Keys and tokens
Keys, secret keys and access tokens management.

Consumer API keys
[Redacted] (API key)
[Redacted] (API secret key)
[Regenerate](#)

Access token & access token secret
[Redacted] (Access token)
[Redacted] (Access token secret)

Read and write (Access level)
[Revoke](#) [Regenerate](#)

Shaping the future of digital business

Esteban Chiner Sanz

Senior Architect

GFT – Data Practice

esteban.chiner@gft.com