

# APIs con Talend

BIG DATA – BUSINESS INTELLIGENCE – MACHINE LEARNING

Enriquecer el flujo de datos con API de datos estadísticos

strate**bi**  
open business intelligence

Agosto 2018



## 1. CASO PRACTICO

### INTRODUCCIÓN

---

El propósito de este documento es demostrar como con un pequeño ejemplo se puede establecer un flujo de datos continuo entre la API Idescat y la herramienta Talend.

El Idescat expone parte de sus datos a través de una colección de API de tipo REST. Además, la API del Idescat ofrece métodos (GET principalmente) para que los programas de terceros puedan obtener información del Instituto de Estadística de Cataluña e integrarla en otros servicios.

Las peticiones sobre la API se componen siempre de un servicio, una versión, una operación, formato y parámetros opcionales.

Los servicios disponibles son:

- **Indicadores al día:** proporciona información básica de una selección de indicadores de Cataluña.
- **Búsqueda de población:** proporciona información de la población por sexo de cualquier entidad territorial de Cataluña.
- **Rectificaciones:** da acceso a la información del registro de *rectificaciones* del sitio web del Idescat.
- **Municipio en cifras:** acceso a información estadística básica comparativa de municipios, comarcas y Cataluña.
- **Calendario de difusión de resultados:** información del calendario del Idescat de difusión de resultados.
- **Días Internacionales:** Api desactivada en 2013.
- **Onomástica:** información estadística de nombres y apellidos de la población y nombres de los recién nacidos.
- **Incrustaciones:** proporciona contenido asociado que puede ser incrustado en páginas de terceros.

Cada uno de los servicios de la API tienen operaciones comunes y/o propias para la solicitud de las peticiones.

Para una mayor referencia sobre todo en términos de parametrización de la API suministramos el siguiente enlace: <https://www.idescat.cat/dev/api/?lang=es>, donde se encontrarán los servicios mencionados y una mayor documentación de la misma.

Por parte de *Talend* se dispondrá de un ejercicio que solicitará datos a la API, los cuales serán tratados y comparados con los documentos *Excel* suministrados.

## FINALIDAD

La finalidad consistirá en el cruce del fichero Excel con las llamadas a la API de *Municipio en cifras* para la obtención de la Comunidad y las Comarcas

Como se puede observar en el documento Excel, se dispone de un campo *Municipi*, el cual corresponde con los municipios de toda Cataluña. Gracias a la disponibilidad del campo se puede cruzar el flujo de datos perteneciente a la API y conseguir una mayor afinidad en el documento Excel mediante los nuevos campos de *Comunitat* y *Comarca*.

Format	Tipus	Data_entrada	Canal_Entrada	Barri	Municipi	Assumptes	clas1	clas2
Comarca	Comarca	2014-2017	Comarca	Les Illes	Can Jussà	Tabla de información de las comarcas de las Islas Baleares, Cataluña y Aragón. Incluye los datos de población, superficie, número de municipios, número de habitantes, número de empresas, número de establecimientos, número de viviendas, número de viviendas con acceso a Internet, número de viviendas con acceso a fibra óptica, número de viviendas con acceso a televisión por cable y número de viviendas con acceso a televisión por satélite.	Comarca	Comarca
Comarca	Comarca	2014-2017	Comarca	Algarbes, Comarques	Can Jussà	Tabla de información de las comarcas de las Islas Baleares, Cataluña y Aragón. Incluye los datos de población, superficie, número de municipios, número de habitantes, número de empresas, número de establecimientos, número de viviendas, número de viviendas con acceso a Internet, número de viviendas con acceso a fibra óptica, número de viviendas con acceso a televisión por cable y número de viviendas con acceso a televisión por satélite.	Comarca	Comarca
Comarca	Comarca	2014-2017	Comarca	Comarques	Can Jussà	Tabla de información de las comarcas de las Islas Baleares, Cataluña y Aragón. Incluye los datos de población, superficie, número de municipios, número de habitantes, número de empresas, número de establecimientos, número de viviendas, número de viviendas con acceso a Internet, número de viviendas con acceso a fibra óptica, número de viviendas con acceso a televisión por cable y número de viviendas con acceso a televisión por satélite.	Comarca	Comarca
Comarca	Comarca	2014-2017	Comarca	Comarques	Can Jussà	Tabla de información de las comarcas de las Islas Baleares, Cataluña y Aragón. Incluye los datos de población, superficie, número de municipios, número de habitantes, número de empresas, número de establecimientos, número de viviendas, número de viviendas con acceso a Internet, número de viviendas con acceso a fibra óptica, número de viviendas con acceso a televisión por cable y número de viviendas con acceso a televisión por satélite.	Comarca	Comarca
Comarca	Comarca	2014-2017	Comarca	Comarques	Can Jussà	Tabla de información de las comarcas de las Islas Baleares, Cataluña y Aragón. Incluye los datos de población, superficie, número de municipios, número de habitantes, número de empresas, número de establecimientos, número de viviendas, número de viviendas con acceso a Internet, número de viviendas con acceso a fibra óptica, número de viviendas con acceso a televisión por cable y número de viviendas con acceso a televisión por satélite.	Comarca	Comarca

Excel

```

-<fitxes version="1.00" lang="ca" o="nodes" p="tipus=cat">
  <v id="09" scheme="ca">Catalunya</v>
</fitxes>

```

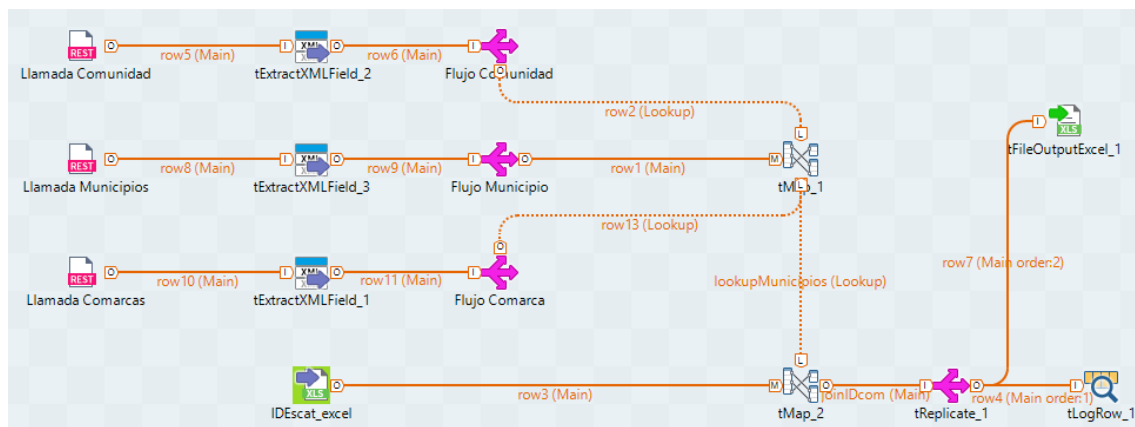
API: Municipio en Cifras – Comunidades

```
<fitxes version="1.00" lang="ca" o="nodes" p="tipus=com">
  <v id="01" scheme="com">Alt Camp</v>
  <v id="02" scheme="com">Alt Empordà</v>
  <v id="03" scheme="com">Alt Penedès</v>
  <v id="04" scheme="com">Alt Urgell</v>
  <v id="05" scheme="com">Alta Ribagorça</v>
  <v id="06" scheme="com">Anoia</v>
  <v id="39" scheme="com">Aran</v>
  <v id="07" scheme="com">Bages</v>
  <v id="08" scheme="com">Baix Camp</v>
  <v id="09" scheme="com">Baix Ebre</v>
```

*API: Municipio en Cifras – Comarcas*

```
<fitxes version="1.00" lang="ca" o="nodes" p="tipus=mun">
  <v id="080018" scheme="mun">Abrera</v>
  <v id="080023" scheme="mun">Aguilar de Segarra</v>
  <v id="080039" scheme="mun">Alella</v>
  <v id="080044" scheme="mun">Alpens</v>
  <v id="080057" scheme="mun">Ametlla del Vallès, l'</v>
  <v id="080060" scheme="mun">Arenys de Mar</v>
  <v id="080076" scheme="mun">Arenys de Munt</v>
  <v id="080082" scheme="mun">Argençola</v>
  <v id="080095" scheme="mun">Argentona</v>
  <v id="080109" scheme="mun">Artés</v>
  <v id="080116" scheme="mun">Avià</v>
  <v id="080121" scheme="mun">Avinyó</v>
  <v id="080137" scheme="mun">Avinyonet del Penedès</v>
```

*API: Municipio en Cifras – Municipios*



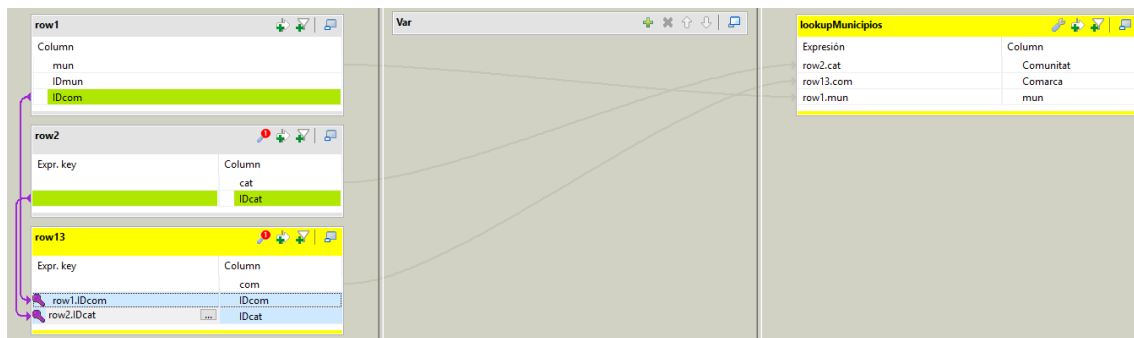
Talend contiene una serie de componentes los cuales permiten trabajar con distintos flujos de datos, archivos y demás. En este caso, los orígenes de datos son un Excel suministrado y las llamadas a la API de Idescat.

Se realizarán un total de tres llamadas para conocer la comunidad, comarcas y municipios disponibles en Cataluña.

A cada una de estas llamadas les corresponde un componente de Talend:

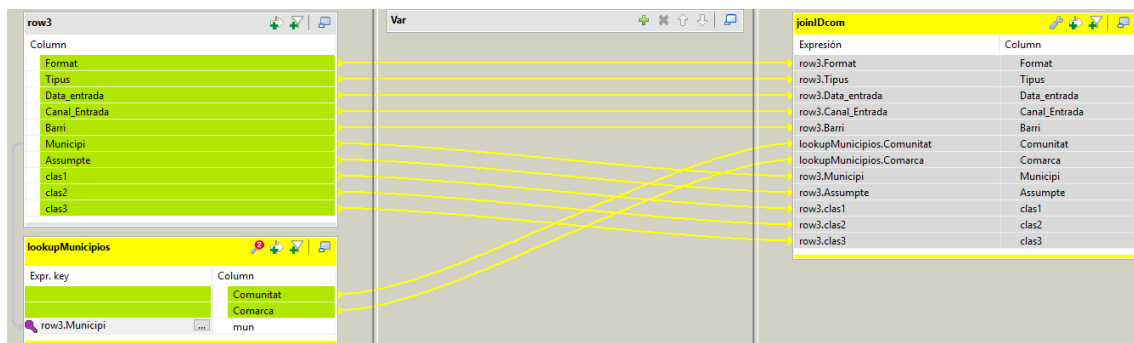


Cada una de las llamadas nos devolverá un flujo de datos distinto, por lo tanto, se transformarán en un único flujo para ser usados contra el documento Excel.



Talend – Conf. componente Tmap

Al disponer del campo *Municipi*, el cual corresponde con los municipios de toda Cataluña. Se puede cruzar con el flujo de datos anterior.



Talend – Conf. componente Tmap2

Para la obtención del flujo de datos Excel hace falta una previa configuración en un componente *tFileInputExcel*, donde se configura la estructura del Excel para obtener los datos correctamente.



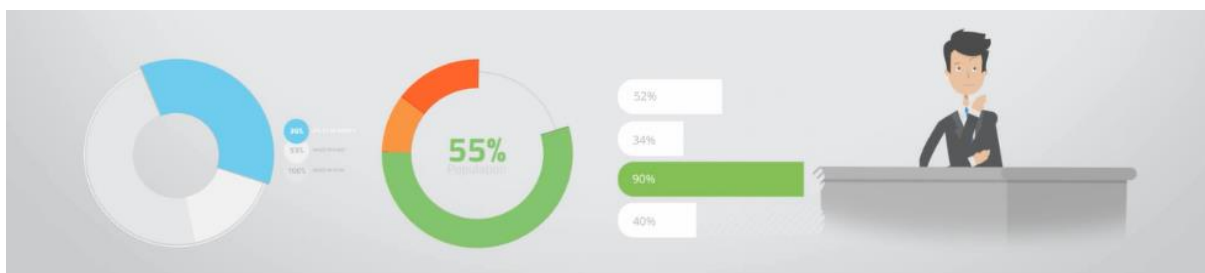
## 2. DESCRIPCIÓN

En Stratebi ofrecemos **gran cantidad de soluciones analíticas** por una compañía de **rápido crecimiento**, innovando en las áreas tecnológicas de mayor desarrollo en la actualidad: **Business Intelligence, Big Data y Social Intelligence**, muchas de ellas, basadas en soluciones **Open Source**.

Además, somos **Partners Certificados en Microsoft PowerBI y Vertica**, con gran número de proyectos con ambas tecnologías

[info@stratebi.com](mailto:info@stratebi.com)

[www.stratebi.com](http://www.stratebi.com)



Desarrollamos nuevas soluciones analíticas basadas en Open Source, para la generación de Cuadros de Mando en tiempo real, con tecnologías IoT para SmartCities, machine learning, etc...





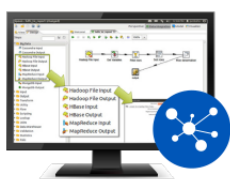
### 3. TECNOLOGÍAS

Recientemente, hemos sido nombrados Partners Certificados de Vertica y Microsoft PowerBI



## 4. EJEMPLOS DE DESARROLLOS ANALYTICS

A continuación se presentan **ejemplos de algunos screenshots** de cuadros de mando diseñados por Stratebi, con el fin de dar a conocer lo que se puede llegar a obtener, así como Demos Online en la web de Stratebi:



Data Ingestion  
Manipulation  
Integration



Enterprise and  
Ad Hoc Reporting



Data Discovery  
Visualization



Predictive  
Analytics

