TOMORROWLAND 2017 BOOM-BELGICA

Bryan Bohórquez

(bryan.bohorquez01@epn.edu.ec)

Santiago Toapanta

(santiago.toapanta@epn.edu.ec)

Jazmin Villamarin

(jazmin.villamarin@epn.edu.ec)

Escuela Politécnica Nacional

Base de Datos Multidimensionales

**RESUMEN:**

En este documento se especificara la arquitectura propuesta para el proyecto mediante el uso de CouchDB (Base de Datos).

Así mismo, se recopilarán diversos datos para posteriormente crear una base de datos unificada de tipo NoSQL analizando los datos obtenidos.

# INTRODUCCIÓN

Sin lugar a dudas, Tomorrowland es el festival de música dance más importante del planeta.

La popularidad de un evento de esta escala se puede sentir desde el principio: las entradas y los packs se agotan en cuestión de minutos cada año.

Tomorrowland es el festival de música electrónica más grande del mundo y ofrece un espectáculo inigualable con los DJs más populares del planeta, un impactante diseño del escenario y espectaculares fuegos artificiales. Todo esto y más hacen de Tomorrowland un festival insuperable.

Desde sus inicios en el 2005, este festival belga se ha convertido en el pináculo mismo de la música dance, presentando únicamente a los mejores DJs y atrayendo a fans de todo el mundo.

Esto nos permite formularnos la pregunta principal:

* + ¿Qué dj tiene más acogida o es favorito en dicho festival?

Por tanto determinaremos la zona donde analizaremos y recolectaremos nuestros datos.

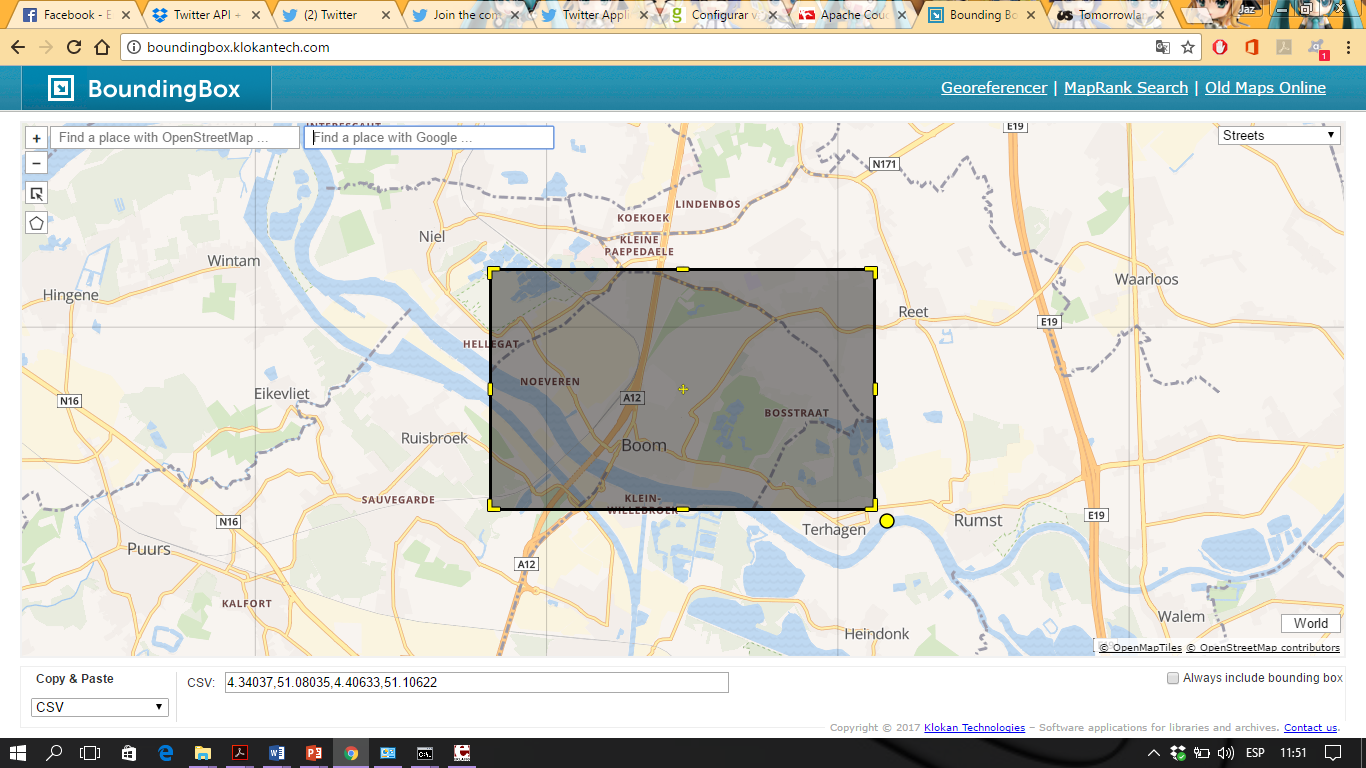


Figura 1. Zona para analizar

# TECNICAS DE RESOLUCION

Arquitectura Data Warehouse

La arquitectura que esta propuesta para el diseño de la BDD NoSQL tiene como base la recopilacion de tweet en CouchDB.

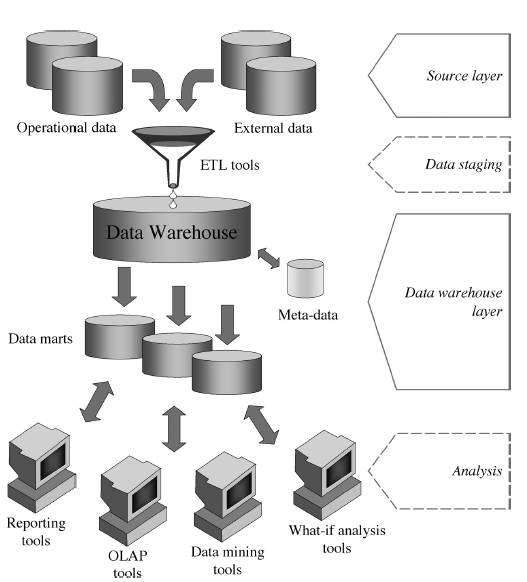


Figura 2. Arquitectura Data Warehouse

Harvesting: El harvesting o cosecha de Tweets es una actividad que no se la puede realizar manualmente, por lo tanto debemos automatizar este procedimiento.

La recolección de tweets se realizó con base a la tendencia de los Djs que se presentaran en el evento Tomorrowland 2017.

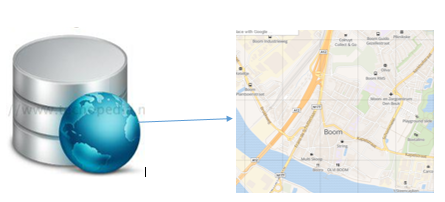


Figura 3. Almacenamiento de datos

Creación de la API para el almacenamiento de datos.

Crearemos la API con el nombre tomorrowland\_epn.

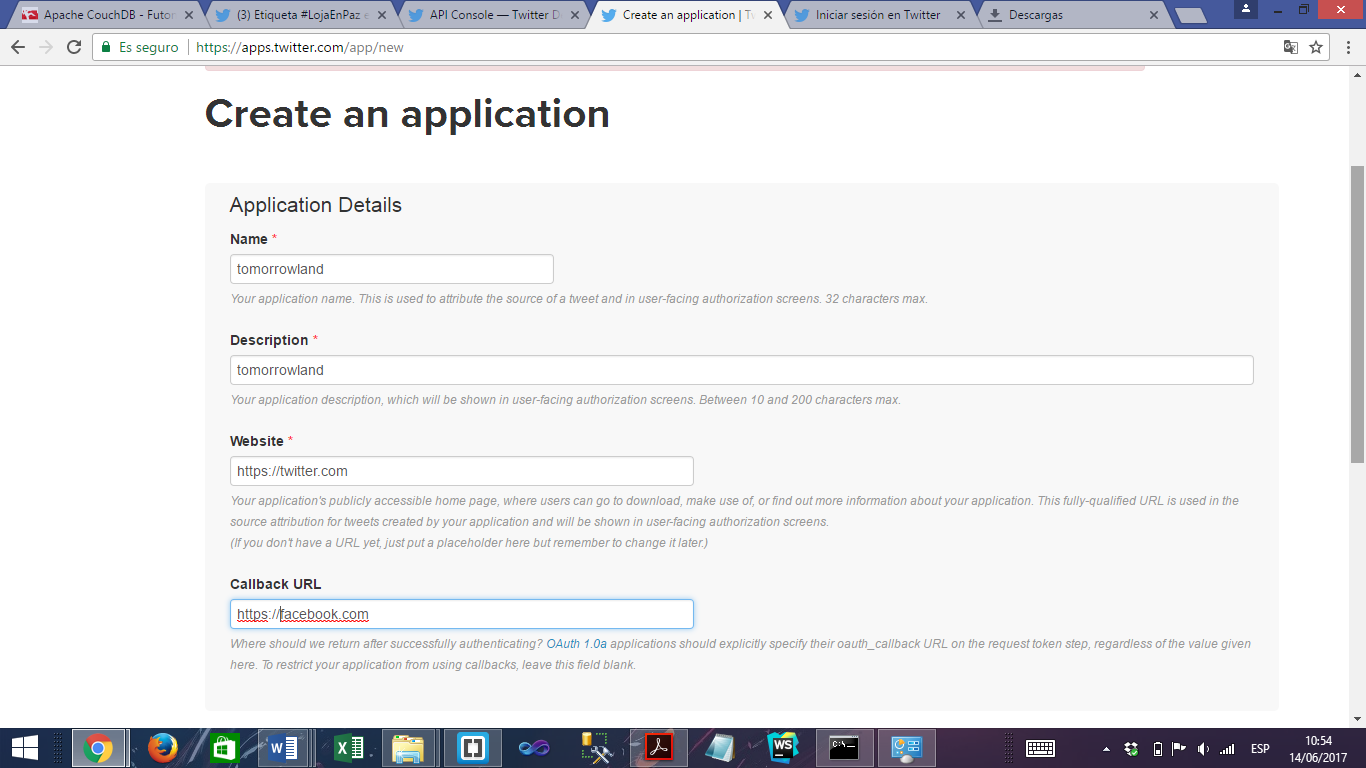


Figura 4. Pasos creación API

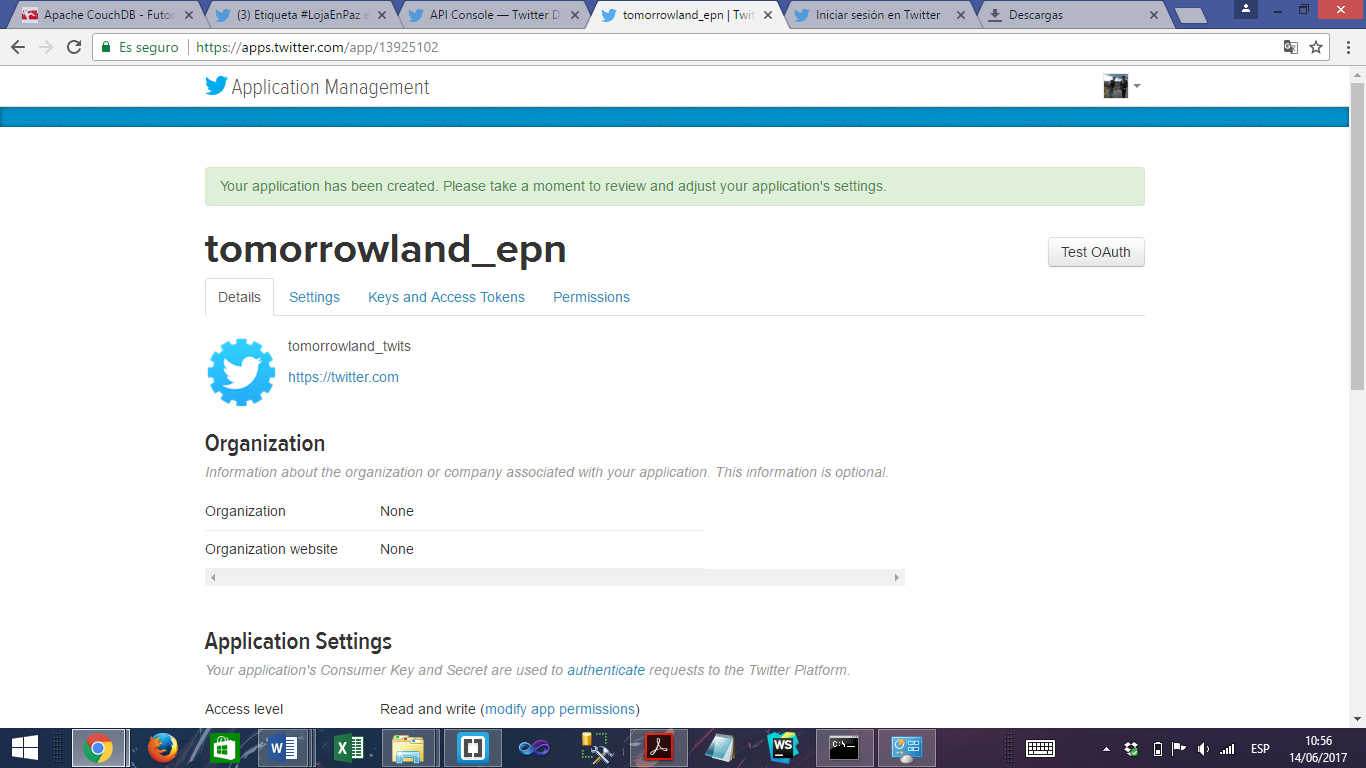


Figura 5. Pasos creación API

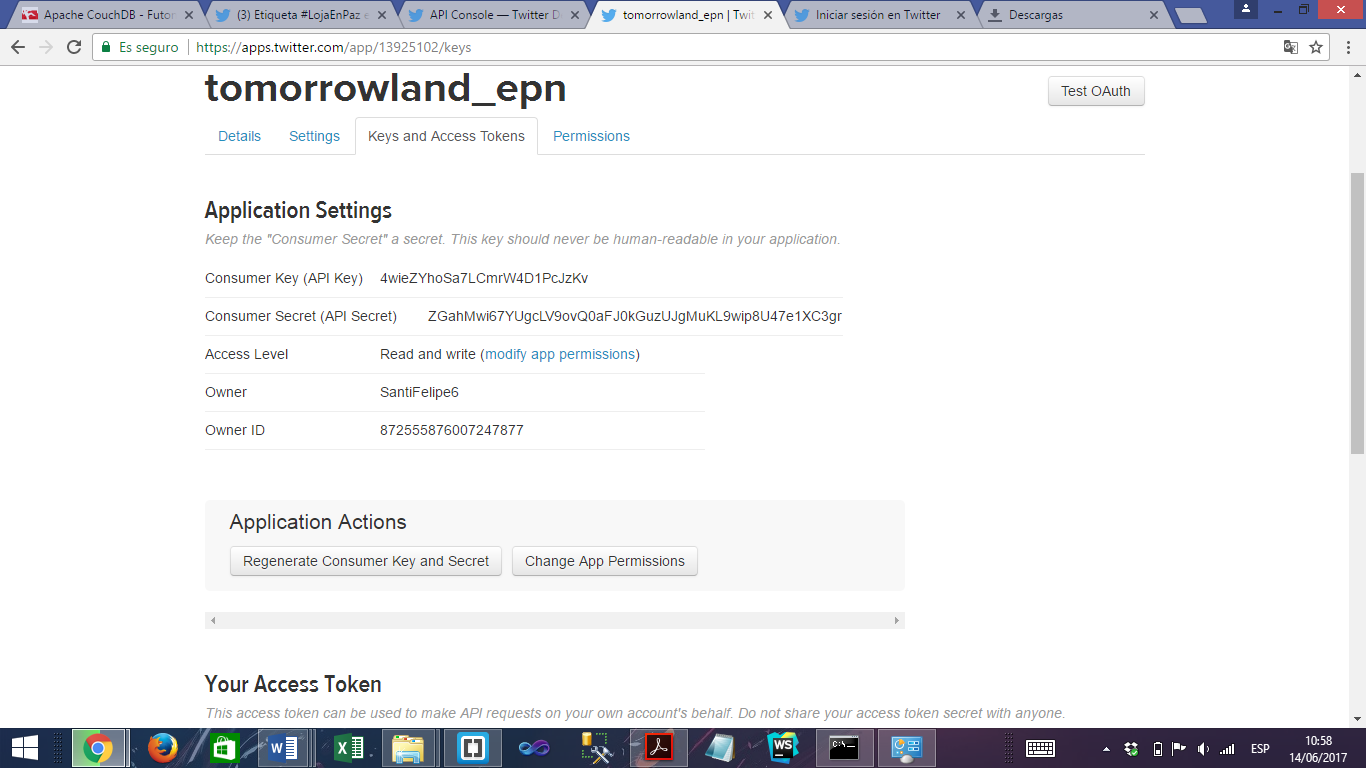


Figura 6. Pasos creación API

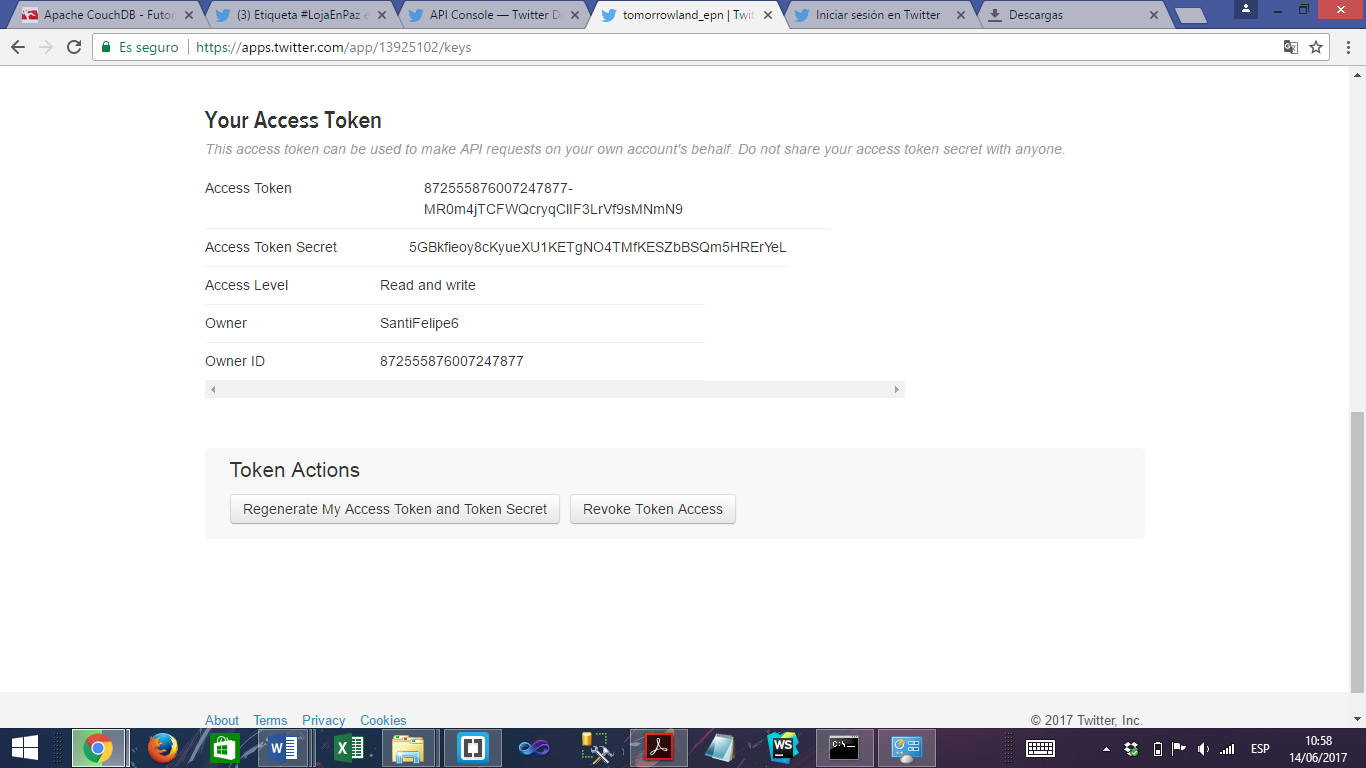


Figura 7. Pasos creación API (credenciales)

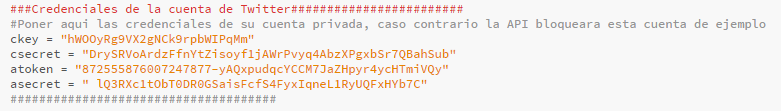
Esta credencial deberá ser introduccida en el archivo de python .py que se encargara de la recoleccion de tweets, a mas de ello en el documento se debe colocar el nombre de la base de datos donde se alojaran los datos y de igual manera la ubicación en un archivo en formato csv del lugar de donde obtendremos los datos.

Para obtener la ubicación y las coordenadas del lugar en csv emplearemo el uso de la Herramienta BOUNDING BOX.



RECOLECCION DE TWEETS

Ingreso de credenciales de la cuenta de Twitter:



Conexión con el servidor CouchDB y crear la respectiva BDD.

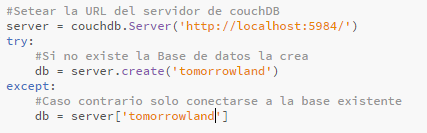


Figura 7. Credenciales

Adicionalmente para ejeutar este tipo d archivos hay que realizar un tipo de coniguraciones por lo cual se itnalan los siguientes paquetes:

* pip install couchdb
* pip install tweepy

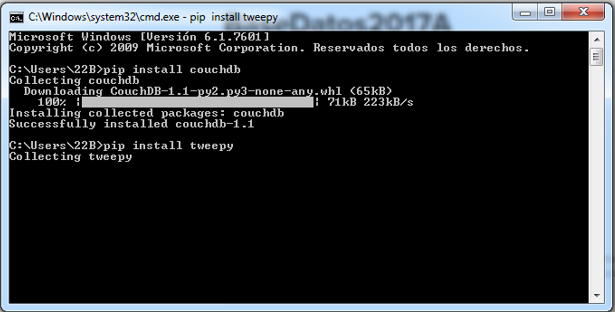


Figura 8. Configuraciones

Recoleccion de datos al ejecutar el archivo de python.

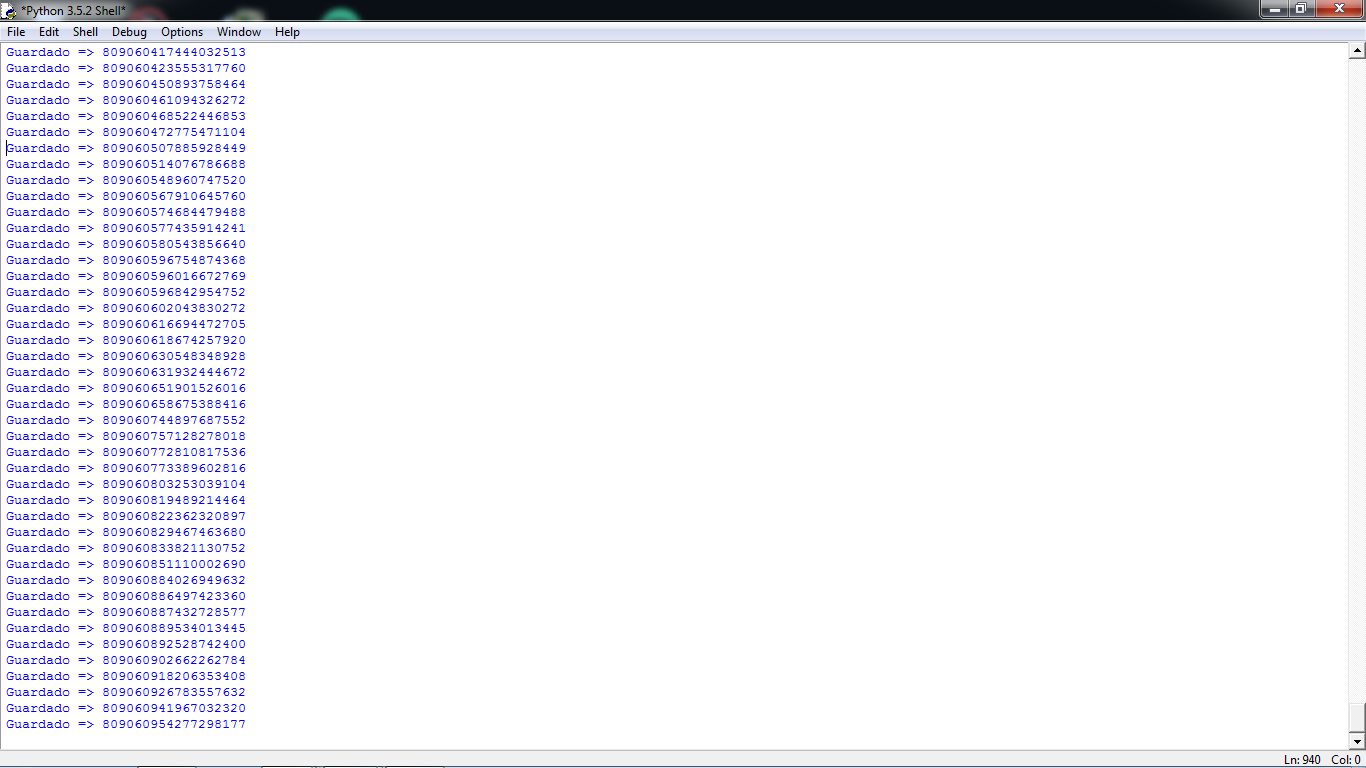


Figura 9. Recopilación de datos

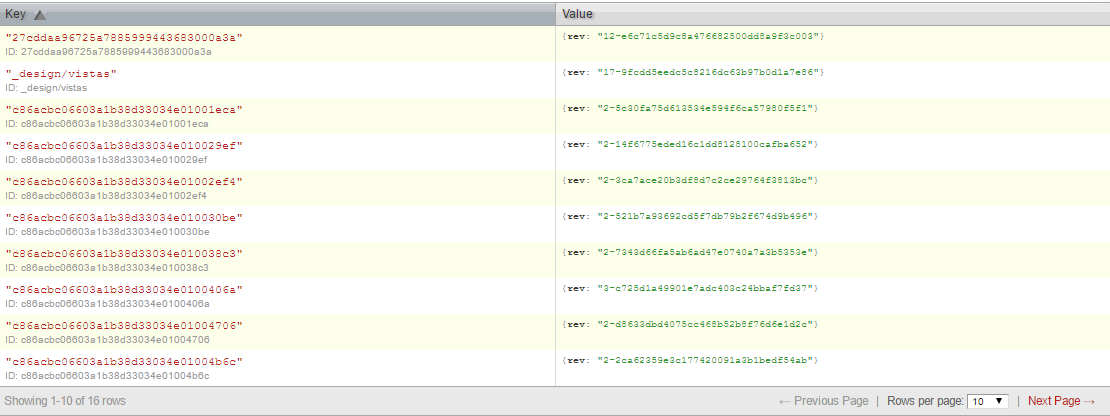


Figura 10. Datos visualizados en CouchDB

# CONCLUSIONES

* En CouchDB, los datos sonuna colección de documentoscompuestos por campos, cuyos valores pueden ser cadenas de texto, números, fechas, listas ordenadas y mapas asociativos.
* Cada documento es identificado por un únicoID es decir un códigoo identificador
* Las APIs REST proporcionan acceso programático a leer y escribir datos de Twitter.

# REFERENCIAS

Hernampérez, R. (2009). CouchDB: Introducción. [online] Rafinguer.blogspot.com. Available at: http://rafinguer.blogspot.com/2009/12/couchdb-introduccion.html

https://dev.twitter.com/rest/public

https://www.festicket.com/es/festivals/tomorrowland/2017/