

In [1]:

```
import re
import pandas as pd
from io import open
import numpy as np
```

In [2]:

```
fichero = open ( "Web_access_log-akumenius.com.txt", "r")

lineas= fichero.readlines()

for i in range(15,30):
    print (lineas[i])
fichero.close()
```

```
localhost 127.0.0.1 - - [23/Feb/2014:03:10:31 +0100] "OPTIONS * HTTP/1.0" 200 - "-" "Apache (internal dummy connection)" VLOG=-
```

```
localhost 127.0.0.1 - - [23/Feb/2014:03:10:31 +0100] "OPTIONS * HTTP/1.0" 200 - "-" "Apache (internal dummy connection)" VLOG=-
```

```
localhost 127.0.0.1 - - [23/Feb/2014:03:10:31 +0100] "OPTIONS * HTTP/1.0" 200 - "-" "Apache (internal dummy connection)" VLOG=-
```

```
localhost 127.0.0.1 - - [23/Feb/2014:03:10:31 +0100] "OPTIONS * HTTP/1.0" 200 - "-" "Apache (internal dummy connection)" VLOG=-
```

```
localhost 127.0.0.1 - - [23/Feb/2014:03:10:31 +0100] "OPTIONS * HTTP/1.0" 200 - "-" "Apache (internal dummy connection)" VLOG=-
```

```
localhost 127.0.0.1 - - [23/Feb/2014:03:10:31 +0100] "OPTIONS * HTTP/1.0" 200 - "-" "Apache (internal dummy connection)" VLOG=-
```

```
localhost 127.0.0.1 - - [23/Feb/2014:03:10:31 +0100] "OPTIONS * HTTP/1.0" 200 - "-" "Apache (internal dummy connection)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:31 +0100] "GET /hoteles-baratos/ofertas-hotel-Club-&-Hotel-Letoonia--en-Fethiye-8460b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:33 +0100] "GET /hoteles-baratos/ofertas-hotel-Metropolis-Hotel-en-Kaunas-210244b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:35 +0100] "GET /hoteles-baratos/ofertas-hotel-Faena-Hotel-Buenos-Aires-en-Downtown-129723b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:38 +0100] "GET /hoteles-baratos/ofertas-hotel-Kensington-Townhouse-London-en-Kensington-54356b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:39 +0100] "GET /destinos-baratos/destinosEstrelles/hoteles-en-Invermere_CANAD%C3%81-con-4-estrellas HTTP/1.1" 200 8811 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:40 +0100] "GET /hoteles-baratos/ofertas-hotel-Howard-Jhonson-hotel-&-Spa--Villa-Gral-Belgrano-en-Villa%20General%20Belgrano-146416b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:42 +0100] "GET /hoteles-baratos/ofertas-hotel-Princesa-Sofia-en-Diagonal%20S.-4013b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:10:45 +0100] "GET /hoteles-baratos/ofertas-hotel-Kfar-Giladi-en-Tiberias-121085b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=-
```

```
In [5]: print (len(lineas)) # miramos la longitud(líneas) del archivo.
```

```
261873
```

Mirando el archivo de manera aleatoria, y en diferentes sitios, podemos diferenciar dos tipos diferentes de líneas.

Una, en que el acceso es desde el *servidor local*. Otra en que es conexión a la web es por transmisión de datos via *www*.

- Rápidamente observamos tres variables: una manera de acceder a la web mediante *www* o *localhost*, una IP y la fecha. Vemos que la manera en que están posicionados en las filas es igual en ambos tipos de filas. El acceso a la web o *Localhost* está separados de la ip mediante un espacio en blanco mientras que la IP y la fecha están separadas de por dos guiones y tres espacios en blanco de la siguiente forma " - - ". Por lo que identificamos 3 variables: Acceso, IP y fecha de acceso
- Tras la variable tiempo observamos dos variables contenidas entre dos pares de comillas. ".....". Vemos que en los registros de servidor local, tenemos dos variables. Una que pone *OPTIONS*, y la versión del *protocolo HTTP*, mientras que en las líneas que acceden via *www*, tenemos el resultado obtenido de lo que aparenta una *búsqueda* y el *protocolo HTTP*.
- *Uno o dos números*: después del protocolo HTTP, tenemos un número de 3 cifras seguidos de un guión en el caso de *localhost*, y de 7 o más cifras en el *WWW*. Pero separados por un espacio de la forma *xxx yyyy*
- Nota. Hasta éste punto las separaciones de datos han sido coherentes. Es decir, hemos encontrado información separada por espacios, o por guiones, en ambos sitios, y las supuestas variables separadas por algún carácter especial o espacio en blanco, tienen su cierta coherencia. En éste punto es el único que las cosas no coinciden. I
- Tras esto tenemos un valor entre comillas, que tal vez es el más difícil de ver, ya que hay pocos registros con datos, como nuestro más abajo. Podemos ver "<http://www.akumenius.com/>" o nada. La llamaremos variable X. Si algo hay en común son las comillas. Por último, vemos otro registro, entre comillas que es el *navegador* con el que se ha accedido.
- Bien, menos por la variable de los números (que no sabemos como opera), y la variable X, el resto parecen fácilmente identificables. A pesar de ello, sí he podido identificar unos patrones de separación de información, de los que me aprovecharé para dejar las líneas preparadas para convertir en *Dataframe*, reconvirtiendo todas estos caracteres comunes en ";

```
In [19]: for i in range(224574,224579):
          print (lineas[i])
          # y podemos ver que hay un registro de una web. También vemos que en el navegador, da más
```

```
www.akumenius.com 37.11.115.207 - - [28/Feb/2014:00:48:12 +0100] "GET /includes/images/uploaded/Slideshow/ESC-RELAX-LANJARON.png HTTP/1.1" 200 42972 "http://www.akumenius.com/" "Mozilla/5.0 (Windows NT 6.0) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/31.0.1650.63 Safari/537.36" VLOG=-
```

```
www.akumenius.com 37.11.115.207 - - [28/Feb/2014:00:48:12 +0100] "GET /modules/raton/view_s/themes/bcoos/images/buscadores/icono-facebook.png HTTP/1.1" 200 2167 "http://www.akumeni
```

```
us.com/" "Mozilla/5.0 (Windows NT 6.0) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/31.0.1650.63 Safari/537.36" VLOG=--
```

```
www.akumenius.com 37.11.115.207 - - [28/Feb/2014:00:48:12 +0100] "GET /modules/raton/view s/themes/bcoos/images/cercle_blau.png HTTP/1.1" 200 3770 "http://www.akumenius.com/" "Mozilla/5.0 (Windows NT 6.0) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/31.0.1650.63 Safari/537.36" VLOG=--
```

```
www.akumenius.com 37.11.115.207 - - [28/Feb/2014:00:48:12 +0100] "GET /includes/images/uploaded/Promociones/p_benidorm.jpg HTTP/1.1" 200 14646 "http://www.akumenius.com/" "Mozilla/5.0 (Windows NT 6.0) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/31.0.1650.63 Safari/537.36" VLOG=--
```

```
www.akumenius.com 37.11.115.207 - - [28/Feb/2014:00:48:12 +0100] "GET /includes/images/uploaded/Promociones/p_granada.jpg HTTP/1.1" 200 18269 "http://www.akumenius.com/" "Mozilla/5.0 (Windows NT 6.0) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/31.0.1650.63 Safari/537.36" VLOG=--
```

```
In [20]: linearr= np.array ( lineas)# convertimos la lista en una array para más adelante vectorizada
```

```
In [22]: # cómo queremos separar todas las referencias por ";", vamos a eliminar todos los puntos y comas del
# documento, sustituyendo "; " por ","
def puntocoma(x):
    x= re.sub( ";"," ",x)
    return x
coma =np. vectorize ( puntocoma)
linearr= coma(linearr)
print (linearr[113]) # tengo la línea 113 localizada con puntos y comas.
```

```
www.akumenius.com 66.249.76.216 - - [23/Feb/2014:03:12:49 +0100] "GET /hoteles-baratos/ofertas-hotel-The-Sukosol-en-Rajatevee-%20Phayathai-60694b-destinos.html HTTP/1.1" 404 3100 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)" VLOG=--
```

```
In [8]: pruebal = linearr[11234].copy()
# usaremos esta línea como campo de pruebas para luego aplicar sobre todo el bloque
```

```
In [23]: # primero, sustituimos los dos primeros espacios en blanco en una ";", de forma que quede el campo de pruebas
pruebal1= re.sub( "\s",";", pruebal,2 )
print (pruebal1)
```

```
www.akumenius.com;81.45.52.213;- - [23/Feb/2014:11:02:43 +0100] "POST /destinos-get HTTP/1.1" 200 26 "http://www.akumenius.com/" "Mozilla/5.0 (Linux; U; Android 4.1.1; es-es; GT-I8190 Build/JRO03H) AppleWebKit/534.30 (KHTML, like Gecko) Version/4.0 Mobile Safari/534.30" VLOG=--
```

```
In [24]: # luego vamos a eliminar los dos primeros espacios en blanco que queden así como los dos que quedan en el campo de pruebas
pruebal2= re.sub( "-","", pruebal1,2)
pruebal2=re.sub("\s","", pruebal2,2)
print(pruebal2)
# de esta manera nos quedan perfectamente separadas 3 variables por ";"
```

```
www.akumenius.com;81.45.52.213;[23/Feb/2014:11:02:43 +0100] "POST /destinos-get HTTP/1.1" 200 26 "http://www.akumenius.com/" "Mozilla/5.0 (Linux; U; Android 4.1.1; es-es; GT-I8190 Build/JRO03H) AppleWebKit/534.30 (KHTML, like Gecko) Version/4.0 Mobile Safari/534.30" VLOG=--
```

G=-

In [25]:

```
# si observamos las siguientes 4 comillas, vemos que casualmente nos separan la variable
# así como el número de la variable X, y esta última del navegador.
prueba13= re.sub("'", ";", prueba12,4)
prueba13= re.sub("'", "", prueba13,1)# luego eliminamos la siguiente comilla porque conviene
# más completamente vacían
prueba13= re.sub("'", ";", prueba13,1)
print(prueba13)
```

```
www.akumenius.com;81.45.52.213;[23/Feb/2014:11:02:43 +0100] ;POST /destinos-get HTTP/1.1;
200 26 ;http://www.akumenius.com/; Mozilla/5.0 (Linux; U; Android 4.1.1; es-es; GT-I8190 B
uild/JRO03H) AppleWebKit/534.30 (KHTML, like Gecko) Version/4.0 Mobile Safari/534.30; VLOG
=-
```

In [28]:

```
# por últimos nos quedaría separar el protocolo HTTP de la búsqueda. Que lo harémos por ú
pru = linearr[113].copy()
def fun(x):
    x1=re.sub( "\s",";", x,2 )
    x2= re.sub( "-", "", x1,2)
    x2=re.sub("\s", "", x2,2)
    x3= re.sub("'", ";", x2,4)
    x3= re.sub("'", "", x3,1)
    x3= re.sub("'", ";", x3,1)
    return x3

fun2 =np. vectorize ( fun)
linearr= fun2(linearr)

print (linearr[22234])
```

```
www.akumenius.com;92.56.114.50;[23/Feb/2014:15:58:04 +0100] ;GET /libraries/anythingSlide
r/images/4a.png HTTP/1.1; 200 2410 ;http://www.akumenius.com/; Mozilla/5.0 (Linux, U, Andr
oid 4.0.4, es-es, SonyEricssonLT22i-o Build/6.1.1.B.1.54) AppleWebKit/534.30 (KHTML, like
Gecko) Version/4.0 Mobile Safari/534.30; VLOG=-
```

In [43]:

```
fichero = open ( "weblog.csv", "w")
fichero.write("Web;IP;Fecha;Busqueda;Numero;origen;Navegador;VLOG\n")
for i in range (0, len(linearr)):
    fichero.write(linearr[i])
fichero.close()
```

In [45]:

```
# creamos el DF
df= pd.read_csv("weblog.csv", sep=";")
df
```

Out[45]:

	Web	IP	Fecha	Busqueda	Numero	origen	Navegador	VLOG
0	localhost	127.0.0.1	[23/Feb/2014:03:10:31 +0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOG
1	localhost	127.0.0.1	[23/Feb/2014:03:10:31 +0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOG
2	localhost	127.0.0.1	[23/Feb/2014:03:10:31 +0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOG

	Web	IP	Fecha	Busqueda	Numero	origen	Navegador	VLI
3	localhost	127.0.0.1	[23/Feb/2014:03:10:31+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC
4	localhost	127.0.0.1	[23/Feb/2014:03:10:31+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC
...	...	...	...	...	...	...	...	...
261868	www.akumenius.com	5.255.253.53	[02/Mar/2014:03:05:39+0100]	GET / HTTP/1.1	200 7528	-	Mozilla/5.0 (compatible, YandexBot/3.0, +http...	VLOC
261869	www.akumenius.com	74.86.158.107	[02/Mar/2014:03:09:52+0100]	HEAD / HTTP/1.1	200 -	-	Mozilla/5.0+ (compatible, UptimeRobot/2.0, htt...	VLOC
261870	localhost	127.0.0.1	[02/Mar/2014:03:10:18+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC
261871	localhost	127.0.0.1	[02/Mar/2014:03:10:18+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC
261872	localhost	127.0.0.1	[02/Mar/2014:03:10:18+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC

261873 rows × 8 columns

In [50]:

df.iloc[223334:223345]

Out[50]:

	Web	IP	Fecha	Busqueda	Numero
223334	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:12+0100]	POST /hotel-list-data HTTP/1.1	20 1509
223335	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:13+0100]	GET /hotel-list HTTP/1.1	20 963
223336	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:13+0100]	GET /modules/raton/views/themes/bcoos/css/2col...	20 1131
223337	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:14+0100]	GET /libraries/jqueryui/js/jquery.ui.button.mi...	20 722

	Web	IP	Fecha	Busqueda	Numer
223338	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:14+0100]	GET /libraries/jqueryui/css/ui-lightness/jquer...	20 247
223339	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:14+0100]	GET /libraries/jqueryui/js/jquery.ui.draggable...	20 1884
223340	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:14+0100]	GET /libraries/jqueryui/js/jquery.ui.dialog.mi...	20 1169
223341	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:15+0100]	GET /libraries/jqueryui/css/ui-lightness/jquer...	20 135
223342	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:15+0100]	GET /modules/raton/views/themes/bcoos/css/butt...	200 4
223343	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:15+0100]	GET /modules/raton/views/themes/bcoos/googleMa...	20 708
223344	www.akumenius.com	88.12.69.210	[28/Feb/2014:00:05:16+0100]	GET /libraries/jqueryui/js/jquery.ui.slider.mi...	20 1010

In [57]:

```
# por último vamos a separa la columna búsqueda en 3.
# podemos ver que hay tres variables. Una que manda un comando de acción que se diferencia
# que alguna opción más. La segunda variable suele ser una búsqueda y por último el protocolo
#separados por espacios. Así que usaremos split para separar y crear nuevas columnas

df.insert(8,"opcion",df["Busqueda"].apply(lambda k: k.split(" ")[0]))
df.insert(9,"busqueda",df["Busqueda"].apply(lambda k: k.split(" ")[1] if len(k.split(" "))
df.tail()
```

Out[57]:

	Web	IP	Fecha	Busqueda	Numero	origen	Navegador	VL
261868	www.akumenius.com	5.255.253.53	[02/Mar/2014:03:05:39+0100]	GET / HTTP/1.1	200 7528	-	Mozilla/5.0 (compatible, YandexBot/3.0, +http...	VLOC
261869	www.akumenius.com	74.86.158.107	[02/Mar/2014:03:09:52+0100]	HEAD / HTTP/1.1	200 -	-	Mozilla/5.0+ (compatible, UptimeRobot/2.0, htt...	VLOC

	Web	IP	Fecha	Busqueda	Numero	origen	Navegador	VLI
261870	localhost	127.0.0.1	[02/Mar/2014:03:10:18+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC
261871	localhost	127.0.0.1	[02/Mar/2014:03:10:18+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC
261872	localhost	127.0.0.1	[02/Mar/2014:03:10:18+0100]	OPTIONS * HTTP/1.0	200 -	-	Apache (internal dummy connection)	VLOC

In [58]:

df.iloc[23334:23345]

	Web	IP	Fecha	Busqueda	Numero	origen		
23334	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:16:59+0100]	GET /destinosCaracteristicas/hoteles-baratos-e...	200 16600	-	Gc	
23335	www.akumenius.com	5.10.83.52	[23/Feb/2014:16:16:59+0100]	GET /destinosEstrelles/hoteles-en-%20(Generica...	200 8800	-	A	
23336	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:17:00+0100]	GET /destinos-baratos/destinosEstrelles/hotele...	200 8858	-	Gc	
23337	www.akumenius.com	180.76.5.28	[23/Feb/2014:16:17:03+0100]	GET /destinos-baratos/destinos-caracteristicas...	200 7196	-	Baic	
23338	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:17:03+0100]	GET /destinos-caracteristicas/hoteles-baratos-...	200 8260	-	Gc	
23339	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:17:08+0100]	GET /destinos-caracteristicas/hoteles-baratos-...	200 9534	-	Gc	
23340	www.akumenius.com	180.76.5.97	[23/Feb/2014:16:17:09+0100]	GET /destinos-baratos/destinos-caracteristicas...	200 8526	-	Baic	
23341	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:17:12+0100]	GET /destinos-caracteristicas/hoteles-baratos-...	200 14719	-	Gc	
23342	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:17:16+0100]	GET /destinos-baratos/destinosEstrelles/hotele...	200 8831	-	Gc	

	Web	IP	Fecha	Busqueda	Numero	origen	
23343	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:17:17+0100]	GET /destinos-caracteristicas/hoteles-baratos-...	200 8321	-	Gc
23344	www.akumenius.com	66.249.76.216	[23/Feb/2014:16:17:21+0100]	GET /destinos-caracteristicas/hoteles-baratos-...	200 7188	-	Gc

In [62]:

```
# eliminamos la columna Búsqueda, y Vlog( que vemos que no aporta nada )

df=df.drop("Busqueda",axis=1)
```

In [63]:

```
df=df.drop("VLOG",axis=1)
```

In [64]:

```
df.iloc[334:345]
```

Out[64]:

	Web	IP	Fecha	Numero	origen	Navegador	opcion	
334	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:21+0100]	404 3100	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/hoteles-ba
335	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:23+0100]	404 3100	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/hoteles-ba
336	www.akumenius.com	180.76.5.200	[23/Feb/2014:03:18:24+0100]	404 3100	-	Mozilla/5.0 (compatible, Baiduspider/2.0, +ht...	GET	/hoteles-ba
337	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:25+0100]	200 8814	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	baratos/destir
338	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:26+0100]	404 3100	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/hoteles-ba
339	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:28+0100]	404 3100	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/hoteles-barat
340	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:30+0100]	404 3100	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/hoteles-ba



	Web	IP	Fecha	Numero	origen	Navegador	opcion	
341	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:33 +0100]	404 3100	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/hoteles-ba
342	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:33 +0100]	200 8815	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	baratos/destir
343	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:35 +0100]	200 16003	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/destinc
344	www.akumenius.com	66.249.76.216	[23/Feb/2014:03:18:37 +0100]	200 8452	-	Mozilla/5.0 (compatible, Googlebot/2.1, +http...	GET	/destinc