OTIC ANALYZER

Alexander Cole Mora

Rafael Herrero Álvarez

Raúl Martín Morales

Daniel Ramos Acosta

Pedro Ramos Rodríguez

Moisés Yanes Carballo

Práctica

Los paquetes utilizados para realizar los scripts han sido:

- **Lsa**: realizar análisis de lenguaje semántico.
- Lattice: crear gráficos multivariable para crear una visualización de los datos.
- **Stringr**: realizar tratamientos con cadenas.
- Tm: minería de texto.
- Wordcloud: creación de un gráfico de nube de palabras.

Script del clima

```
library (tm)
library (wordcloud)
library (lsa)
#Cargamos ficheros
raiz <- setwd("~/")</pre>
texto1 <-
readLines(paste(raiz,"/proyecto-final-TIO/datos/londres/Clima1.txt"
,sep = ""),encoding="UTF-8")
texto1 = iconv(texto1, to="ASCII//TRANSLIT")
texto2 <-
readLines(paste(raiz,"/proyecto-final-TIO/datos/londres/Clima2.txt"
,sep = ""),encoding="UTF-8")
texto2 = iconv(texto2, to="ASCII//TRANSLIT")
#Unimos todos los ficheros bajo una única lista y creamos el corpus
docs <- list(texto1,texto2)</pre>
corpus <- Corpus(VectorSource(docs))</pre>
#Limpiamos el conjunto de los documentos de numeros, espacios en
blanco, etc...
d <- tm map(corpus, content transformer(tolower)) #Lo convierte a</pre>
minúsculas
```

```
d <- tm map(d, removePunctuation) #Elimina signos de puntuación</pre>
d <- tm map(d, removeNumbers) #Elimina los números</pre>
d <- tm map(d, stripWhitespace) #Elimina los espacios en blanco</pre>
d <- tm map(d, removeWords,</pre>
c("london","weather","climate","temperatures","temperature","year
", "months",
"usually", "although", "less", "day", "city", "cities", "likely", "capit
al", "normally",
"always", "however", "sometimes", "throughout", "idea", "common", stopw
ords("english"))) #Elimina palabras vacías en inglés
#Creamos la matriz de términos
tdm <- TermDocumentMatrix(d)</pre>
m = as.matrix(tdm)
# conteo de palabras en orden decreciente
wf <- sort(rowSums(m),decreasing=TRUE)</pre>
# crea un data frame con las palabras y sus frecuencias
dm <- data.frame(word = names(wf), freq=wf)</pre>
# Frecuencia minima igual a 20
wordcloud(dm$word, dm$freq, min.freq = 1,
          max.words=50, random.order=FALSE,
colors=brewer.pal(8,"Dark2"))
```

Script de lugares

```
library (tm)
library (wordcloud)
library (lsa)
#Cargamos ficheros
raiz <- setwd("~/")
texto1 <-
readLines(paste(raiz,"/proyecto-final-TIO/datos/londres/Lugares1.tx
t",sep = ""),encoding="UTF-8")
texto1 = iconv(texto1, to="ASCII//TRANSLIT")
texto2 <-
readLines(paste(raiz, "/proyecto-final-TIO/datos/londres/Lugares2.tx
t", sep = ""), encoding="UTF-8")
texto2 = iconv(texto2, to="ASCII//TRANSLIT")
texto3 <-
readLines(paste(raiz, "/proyecto-final-TIO/datos/londres/Lugares3.tx
t", sep = ""), encoding="UTF-8")
texto3 = iconv(texto2, to="ASCII//TRANSLIT")
#Unimos todos los ficheros bajo una única lista y creamos el corpus
docs <- list(texto1,texto2,texto3)</pre>
corpus <- Corpus(VectorSource(docs))</pre>
#Limpiamos el conjunto de los documentos de numeros, espacios en
blanco, etc...
d <- tm map(corpus, content transformer(tolower)) #Lo convierte a</pre>
minúsculas
```

```
d <- tm map(d, removePunctuation) #Elimina signos de puntuación</pre>
d <- tm map(d, removeNumbers) #Elimina los números</pre>
d <- tm map(d, stripWhitespace) #Elimina los espacios en blanco</pre>
d<-tm map(d, removeWords,</pre>
c("london", "londons", "abbey", "cast", "count", "end", "fine", "inside
","mall","new","plan","rich","sir","stay","top","war","aboard","
time", "also", "known", "see", "pauls", "two", "take", "big", "home", sto
pwords("english"))) #Elimina palabras vacías en inglés
#Creamos la matriz de términos
tdm <- TermDocumentMatrix(d)</pre>
m = as.matrix(tdm)
# conteo de palabras en orden decreciente
wf <- sort(rowSums(m),decreasing=TRUE)</pre>
# crea un data frame con las palabras y sus frecuencias
dm <- data.frame(word = names(wf), freq=wf)</pre>
# Frecuencia minima igual a 20
counts <- table (dm$word)</pre>
wordcloud(dm$word, dm$freq, min.freq = 1,
          max.words=50, random.order=FALSE,
colors=brewer.pal(8, "Dark2"))
```

Script de procedencia

```
# Poner la ruta de donde está el repo
rutaBase <- "D:\\Documentos\\Repos\\Clases\\tio\\proyecto"</pre>
rutaCSV <- paste(rutaBase, "/datos/nuevayork/visitantes.csv", sep="")</pre>
rutaOutput <- paste(rutaBase, "/datos/nuevayork/Graficos/", sep="")</pre>
data <- read.csv(rutaCSV, header=TRUE)</pre>
for(i in 2:ncol(data)) {
  visitantes <- data[, c(1, i)]</pre>
  png(filename=paste(rutaOutput, i-1, ".png"),
      units="px",
      width=1024,
      height=700,
      pointsize=12,
      res=72)
  plot(visitantes,type="o")
  dev.off()
```

Canarias. Clima

- Temperatura
- Precipitaciones
- Clima
- Tiempo
- Zonas
- Mar
- Costa
- Lluvias
- Alisio
- Vientos

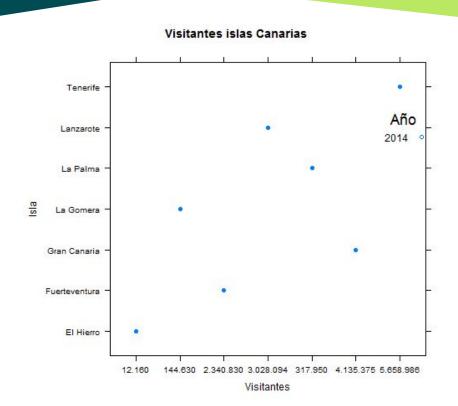
```
perturbaciones
                                                       climaticos archipiaclago
                              geograficas humedad dato vegetacian condiciones relieve faro dato fuerteventura parte grados tenerife temperaturas paases
                       insolacian tipo consultada formacian pueden palisio vientos valores temporada antalyaisla
                             ciudades tempe
edimburgo recoge fuente marser toronto importante ciudad fuente marser toronto importante ciudad fuente marser toronto anual orden tiempo forma laguna basica laguna verano verano delgada eropuerto precipitaciones mensual meses
                                        enero canaria mensual mese
tambiacn frecuencia general
invasiones madrid rojo
superficie maxima
```

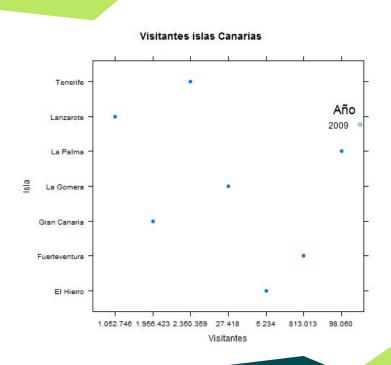
Canarias. Lugares

- Tenerife
- Teide
- Playas
- Paque
- Iglesias
- Capital
- Adeje
- Garachico

```
garachico candelaria comaon visitar ofrece adeje grandes tel plaza plaza
                                                                                                                                                                                                                                                                                                                                                                                                                  gigantes
roque museosan is a canaria con espacio playas parque casa santa espacio playas parque casa santa espacio playas parque casa santa espacio rural laguna palma espacio rural laguna palma espacio rural lugar naturaleza municipio casi municipio casi informacian cristabal
                                                                                                                                                                                                                                                                      wwwlaguiadegrancanariacom
```

Canarias. Procedencia





Londres. Clima

- Precipitation
- Rain
- Rainy
- Cloudy
- Umbrella
- December
- August
- Summers
- IClimate

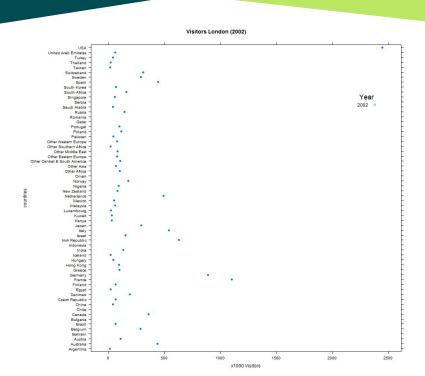
```
visiting summers
   iclimate experience
 ⊆generally of august twice
                  Dbuildings
settling december reputation
  precipitations typical
```

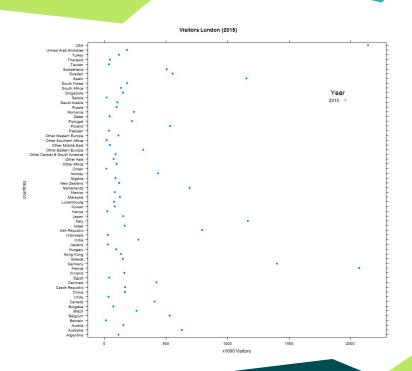
Londres. Lugares

- Tower
- Palace
- Buckingham
- Tower
- Gallery
- Gardens
- Square
- Cathedral



Londres. Procedencia





Nueva York. Clima

- Palabras Clave:
 - Clima
 - Niebla
 - Aire
 - Tormenta
 - Nubes



Nueva York. Lugares

- Central Park
- Broadway
- Manhattan
- Museum
- Avenue
- Madison
- Brooklyn
- Square

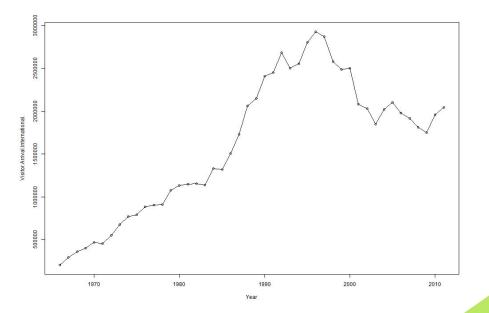
```
building famous brooklyn

music artrestaurants
times club restaurant first
fifthbar of street west
dining street west
dining park avenue

park avenu
```

Nueva York. Procedencia

- Crecimiento hasta 1995
- Luego empezó a descender
- ¿Posible entrada de la crisis?



Modelo de Negocio

El objetivo es obtener información útil de diferentes lugares para venderla a diferentes empresas y organizaciones:

- TourOperadores
- Compañías de vuelo, barcos
- Alquileres de coches
- Agencia de Viajes
- Gobiernos

FIN

Alexander Cole Mora

Daniel Ramos Acosta

Rafael Herrero Álvarez

Pedro Ramos Rodríguez

Raúl Martín Morales

Moisés Yanes Carballo