Googling trends in conservation biology using R
Appendix 1

Raphael Proulx[1], Philippe Massicotte[1], and Marc Pépino[1]

[1]Department of environmental sciences, UQTR

November 1, 2013

1 Introduction

We developed a R package named *GTrendsR* that provides an interface for retrieving and displaying the information returned online by Google Trends in the R console. To install the package (hosted on CRAN repository), simply type the following command in the R console.

```
library(devtools)
install_bitbucket("GTrendsR", "persican")
library("GTrendsR")
```

Basically, the package has three functions that need to be invoked in a specific order by the user.

1. *ch = gConnect(usr, psw)* – Create a connection with Google Trends service.
2. *data = gTrends(ch, geo = location, query = query)* – Perform a **query** at the specified location **geo**.
3. *gTrendsMap(data)* – Plot maps showing results of the query and the regions and cities levels.

2 Connecting to Google Trends service

In order to perform Google Trends queries, the user needs to own a free Google account. The connection to Google Trends service is then established using account information as follow.

```
usr = "username@gmail.com"
psw = "yourpassword"
ch = gConnect(usr, psw)
```

3 Perform a query

A Google Trends query is simply made using the *gTrends* function. The user need to provide three parameters. First, a valid Google service connection **ch** returned by the function *gConnect*. The second parameter **geo** is the geographic location where the query will be performed. This argument should be a string of two characters. To obtain a list of valid country codes supported by Google Trends, the user can simply type:

```
data(countries)
```

The third and last parameter **query** is a string containing the keywords to use. For example, the following query will search for **NHL** keyword in **Canada**.

	CODE	COUNTRY
1	AF	Afghanistan
2	AL	Albania
3	DZ	Algeria
4	AS	American Samoa
5	AD	Andorra
6	AO	Angola

Table 1: Example of country codes.

```
location = "CA"
query = "NHL"
data = gTrends(ch, geo = location, query = query)
```

3.1 Structure of returned data

```
## [1] "Google login successful!"
```

The data returned by *gTrends* is a list of seven *dataframe* described as follow.

`SearchInfo` – Contains the date and the keyword(s) used for the query.

```
data$SearchInfo

##            SearchTerms
## 1 Search interest: nhl
```

`Weeks` – Contains the RAW results of the query.

```
head(data$WeeklyHits)

##                     sDates Hits      Dates Months Years Days
## 1 2004-01-04 - 2004-01-10   25 2004-01-04     01  2004   04
## 2 2004-01-11 - 2004-01-17   25 2004-01-11     01  2004   11
## 3 2004-01-18 - 2004-01-24   24 2004-01-18     01  2004   18
## 4 2004-01-25 - 2004-01-31   23 2004-01-25     01  2004   25
## 5 2004-02-01 - 2004-02-07   23 2004-02-01     02  2004   01
## 6 2004-02-08 - 2004-02-14   24 2004-02-08     02  2004   08
```

`Regions` – Contains the hits for the top 15 regions of the country where the query has been performed.

```
data$Regions

##                   Regions Hits
## 1                 Alberta  100
## 2       British Columbia   95
## 3                 Nunavut   94
## 4                Manitoba   89
## 5                 Ontario   82
## 6             Nova Scotia   74
## 7                  Quebec   72
## 8  Northwest Territories   72
## 9            Newfoundland   71
## 10           Saskatchewan   71
## 11  Prince Edward Island   68
## 12          New Brunswick   65
## 13       Yukon Territory   58
```

`Cities` – Same as regions but for cities.

```
data$Cities
```

```
##        Cities Hits
## 1    Edmonton  100
## 2     Calgary   93
## 3    Winnipeg   88
## 4   Vancouver   87
## 5      Ottawa   84
## 6    Montreal   79
## 7      Surrey   74
## 8     Toronto   73
## 9     Halifax   69
## 10  St John's   66
## 11     London   64
## 12   Victoria   61
## 13  Saskatoon   60
## 14   Hamilton   59
## 15 Burlington   57
```

`TopSearches` – Top searches related to the keyword(s) used for the query.

```
head(data$TopSearches)
```

```
##        Keywords Hits
## 1        hockey  100
## 2     nhl draft   75
## 3 nhl standings   75
## 4       tsn nhl   65
## 5           tsn   65
## 6    nhl scores   65
```

`MonthlyHits` – Contains the normalized results of the query rescaled by dividing the number of search hits obtained for a given country, or region, by the maximum number of hits obtained over the specified period. This temporal serie match the visual display provided upon a query on Google Trends web site.

```
head(data$MonthlyHits)
```

```
##   Months Years  Hits      Dates
## 1     01  2004 30.89 2004-01-01
## 2     02  2004 40.13 2004-02-01
## 3     03  2004 38.54 2004-03-01
## 4     04  2004 60.19 2004-04-01
## 5     05  2004 38.22 2004-05-01
## 6     06  2004 19.43 2004-06-01
```

4 Plot the trends and the distribution maps

A query to Google Trends can be rapidly visualized using the *gTrendsMap* function as follow (see Figure 1):

```
gTrendsMap(data)
```

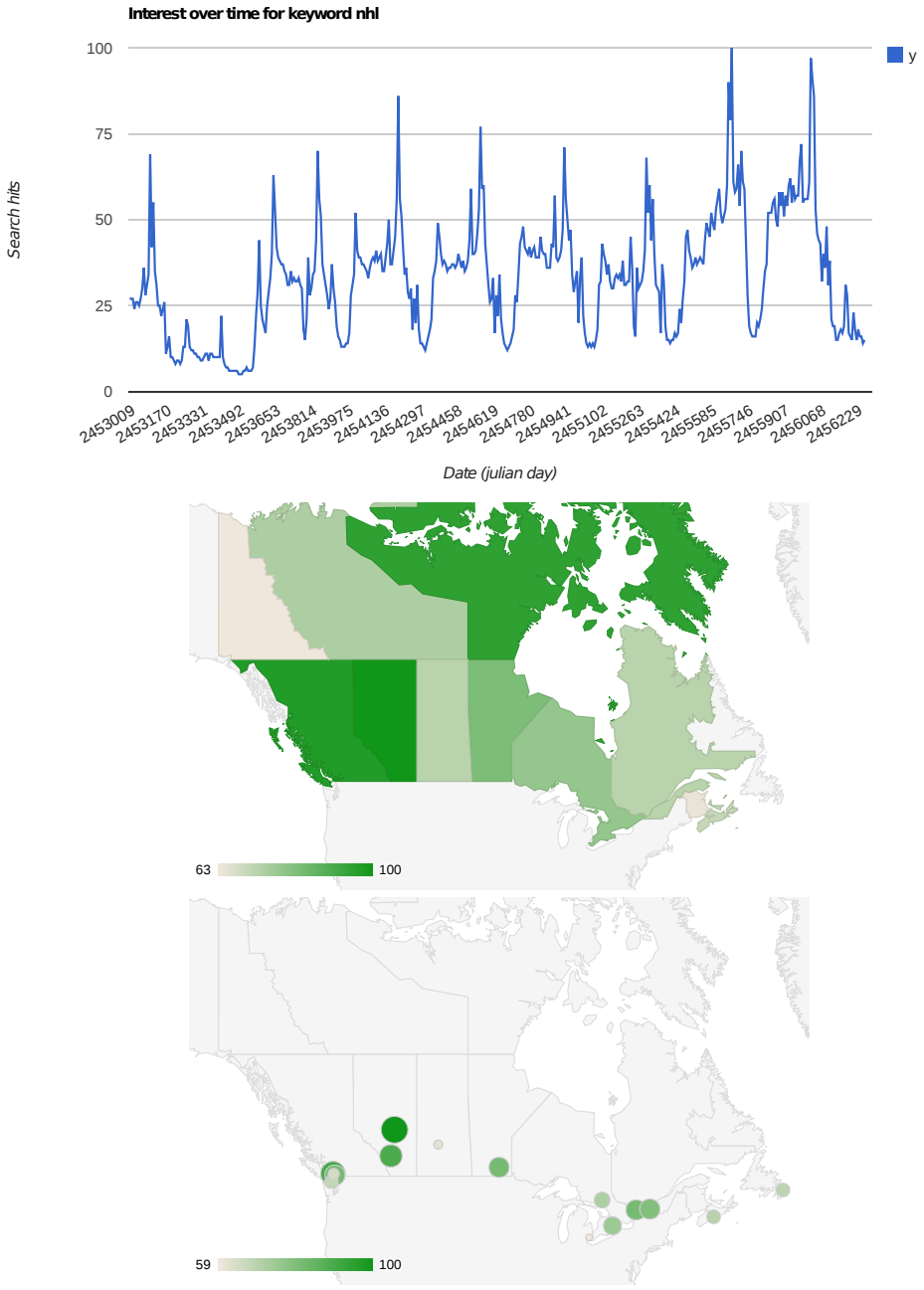

Figure 1: Distribution map for search term *NHL* in *Canada*.