

Ibbe-shiny

un exemple de serveur *shiny*
fait-maison, au service
de la recherche locale

Aurélie Siberchicot
Bruno Spataro
Stéphane Delmotte

Pôle Informatique
LBBE
Université Lyon 1

6^{èmes} Rencontres R
28/30 juin 2017 - Anglet

Ibbe-shiny

un exemple de serveur *shiny*
fait-maison, au service
de la recherche locale

Aurélie Siberchicot
Bruno Spataro
Stéphane Delmotte

Pôle Informatique
LBBE
Université Lyon 1

6^{èmes} Rencontres R
28/30 juin 2017 - Anglet

Contexte & problématique scientifiques

- diffuser des outils et des documents librement et gratuitement
- rendre reproductibles des résultats de recherche
- se soumettre aux contraintes de mise à disposition des journaux

Quelques solutions autour de R

- construire un package *R* ou *BioConductor*
- héberger son application *shiny* sur shinyapps.io
- déployer un serveur *shiny*
- réaliser une page ou un site web
- ...

Chez nous, au LBBE

- une centaine de chercheur(se)s et enseignant(e)s-chercheur(se)s et autant de thésard(e)s, stagiaires, post-docs
- des domaines de recherche variés : écologie, biométrie, génomique, génétique, épidémiologie, phylogénie, ...
- deux thématiques transversales : la biologie évolutionne & l'approche méthodologique par la modélisation mathématique et informatique

Chez nous, au LBBE

- un pôle informatique : 8 techniciens et ingénieur(e)s
- des compétences en développement (site web, progiciels), en optimisation et parallélisation de codes, en bases de données, en bioinformatique, ...
- des ressources informatiques locales : des serveurs, un cluster de calcul (30 nœuds, 1224 cœurs), des systèmes de stockage, ...

Notre cahier des charges

- bénéficier d'un système gratuit
- héberger les sources et les projets localement
- maintenir une proximité avec les chercheurs du LBBE
- proposer de la flexibilité pour répondre aux diverses utilisations
- assurer la sécurité et la sauvegarde
- diffuser des applications *shiny* et des documents *RMarkdown*

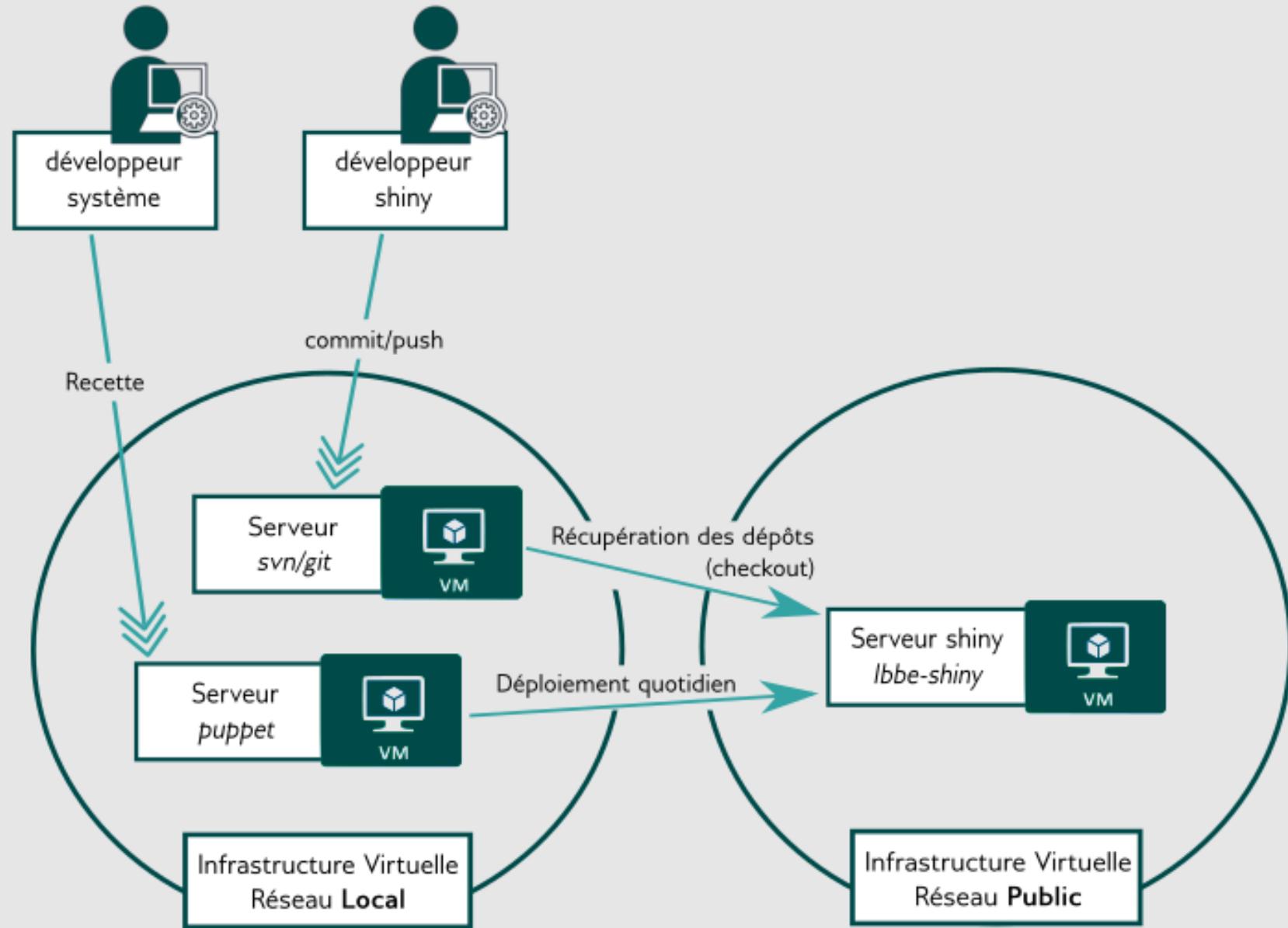
Notre solution sur mesure : *Ibbe-shiny*

- une machine virtuelle (*VMware*, *4 CPUs*, *8Gb de RAM*, *50Gb de disque*)
- un serveur *shiny* (*v1.5.3.838*)
- une recette (*puppet*) élaborée pour une mise à jour quotidienne du système (mise à jour de *R*, des packages, des modules de la VM, checkout des dépôts *Subversion*, ...)
- chaque projet est versionné (*Subversion*)

Notre solution sur mesure : *Ibbe-shiny*

- espace dédié aux applications *shiny*, un autre pour la recherche reproductible
- chaque projet contient des métadonnées :
 - un fichier *install.R* contenant les commandes *R* pour installer les packages *R* nécessaires
 - un fichier *intro.txt* contenant une phrase explicative du projet

Notre solution sur mesure : *Ibbe-shiny*



Ibbe-shiny : avantages

- flexibilité
- maintenance locale
- obligation de versionner les projets
- les différents serveurs sont sur des hôtes différents
- l'intégration du serveur *shiny* dans une VM permet de modifier les ressources allouées à celle-ci

Ibbe-shiny : inconvénients

- instabilité du serveur *shiny*
- capacité de charge limitée
- obligation de versionner les projets

L'avenir de *Ibbe-shiny*

- améliorer l'organisation des projets sur l'interface
- proposer le versionnage par *git*
- exploiter au maximum l'automatisation des installations, des mises à jour, de la maintenance
- étendre l'utilisation de *Ibbe-shiny* au sein du LBBE
- faire « valider » cet outil par les journaux scientifiques

<http://lbe-shiny.univ-lyon1.fr/>

aurelie.siberchicot@univ-lyon1.fr



Lightning Talk

29 JUIN 2017 - 7EMES RENCONTRES R, ANGLET

Analyse des débats sur le secret bancaire au Parlement suisse avec R

Dr. Marion DEVILLE

Dr. Roy GAVA

Dr. Yannick ROCHAT

Emmanuel ROUSSEAU

Pierre-Alexandre FONTA

Université de Genève

Université de Lausanne

Lightning Talk

29 JUIN 2017 - 7EMES RENCONTRES R, ANGLET

Analyse des débats sur le secret bancaire au Parlement suisse avec R

Dr. Marion DEVILLE

Dr. Roy GAVA

Dr. Yannick ROCHAT

Emmanuel ROUSSEAU

Pierre-Alexandre FONTA

Université de Genève

Université de Lausanne





AMTLICHES BULLETIN – BULLETIN OFFICIEL

Ständerat • Wintersession 2014 • Erste Sitzung • 24.11.14 • 16h15 • 10.450
Conseil des Etats • Session d'hiver 2014 • Première séance • 24.11.14 • 16h15 • 10.450



10.450

**Parlamentarische Initiative
FDP-Liberale Fraktion.
Den Verkauf von Bankkundendaten
hart bestrafen**

**Initiative parlementaire
groupe libéral-radical.
Réprimer durement la vente
de données bancaires**

Zweitrat – Deuxième Conseil

CHRONOLOGIE

NATIONALRAT/CONSEIL NATIONAL 27.09.13 (FRIST - DÉLAI)

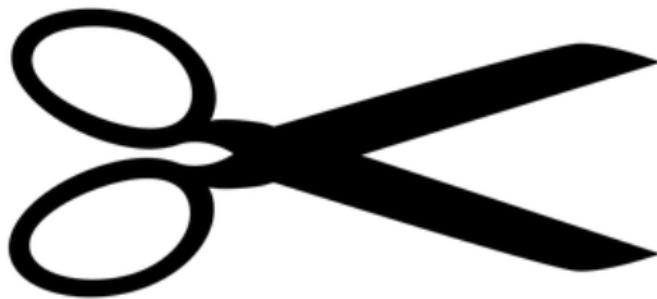
NATIONALRAT/CONSEIL NATIONAL 16.09.14 (ERSTRAT - PREMIER CONSEIL)

STÄNDERAT/CONSEIL DES ETATS 24.11.14 (ZWEITRAT - DEUXIÈME CONSEIL)

NATIONALRAT/CONSEIL NATIONAL 12.12.14 (SCHLUSSABSTIMMUNG)

Etude
Corpus
Packages
Visualisations

Sélection
Extraction interventions
Extraction acteurs
Compositions matrices





Etude
Corpus
Packages
Visualisations

Sélection
Extraction interventions
Extraction acteurs
Compositions matrices





The spnet package Plotting (social) networks on maps

CRAN 0.9.1-0 downloads 6436

Social networks analysis has received special attention over the past decade, and a lot of tools for manipulating and rendering social networks have emerged. In several situations a social network is associated with a spatial dimension, and behaviors observed within the network cannot be interpreted without taking into account the location of each of its nodes regarding to the other nodes. This is the case, for example, when studying inflows/outflows between cities or companies, or when studying people debating in a room. Based on the [sp](#) package, which provides efficient classes for storing spatial data and methods for handling and rendering them, the [spnet](#) package aims at facilitating the rendering of (social) networks on maps. Furthermore, fixing network node positions allows to more easily monitor time-varying networks and observe how connections and flows evolve over time. In the future the package will be also designed to be used together with the [Rsocialdata](#) package for which tools for preparing network data are in development.

Authors: Emmanuel Rousseaux, Marion Deville and Gilbert Ritschard.

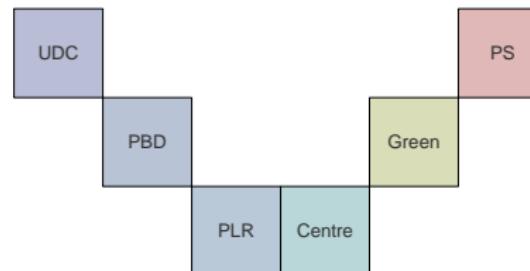
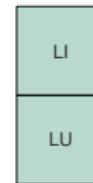
Availability:

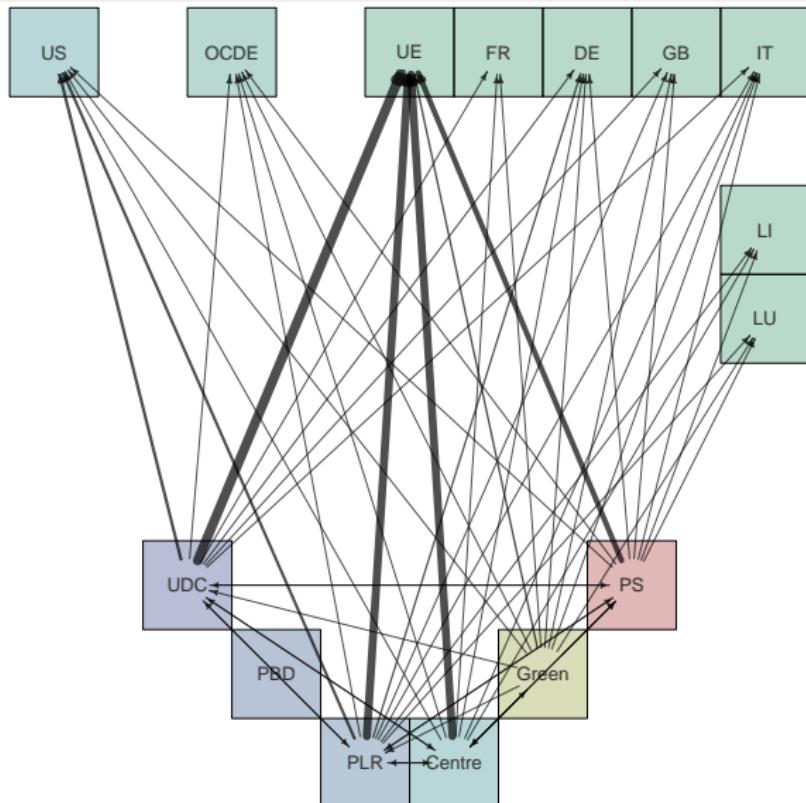
- Official released available from the [CRAN](#).
Installation: `install.packages("spnet")`
- Development version available from the [spnet R-Forge repository](#).
Installation: `install.packages("spnet", repos="http://R-Forge.R-project.org", type = "source")`

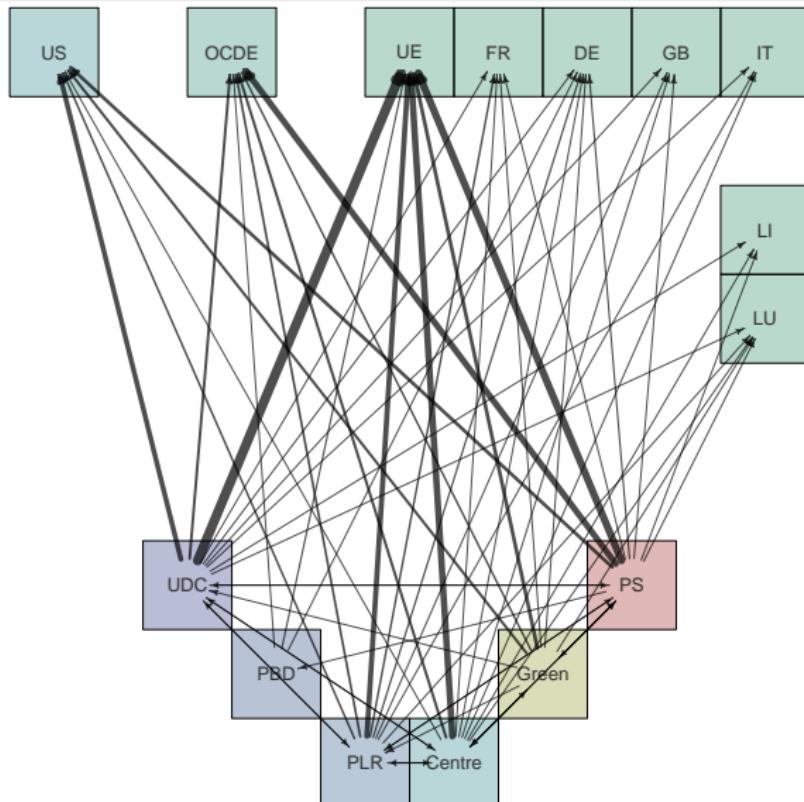
Documentation:

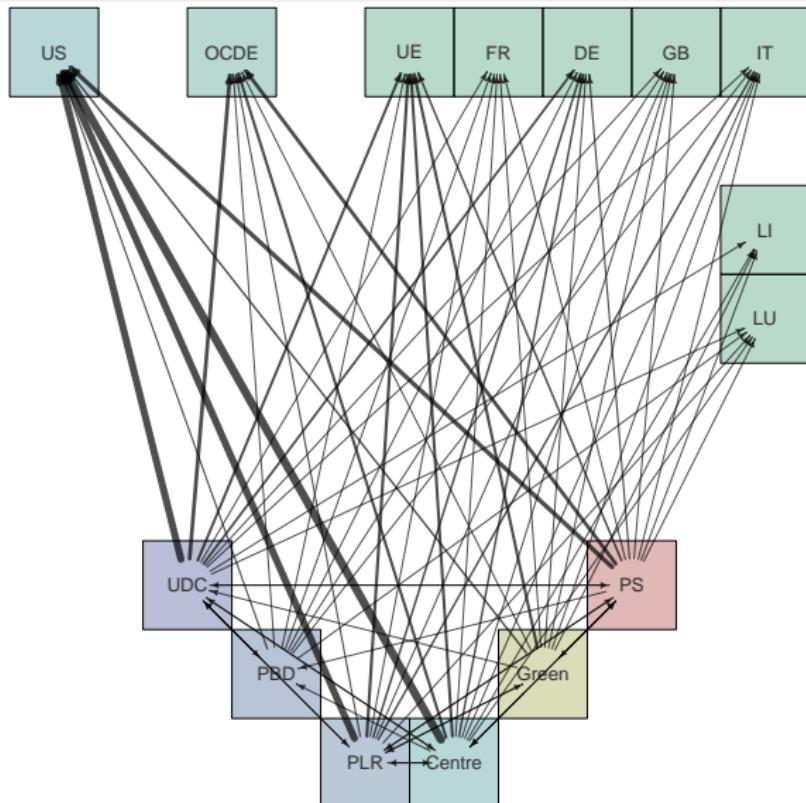
- Package vignette: *Overview of the spnet R package* ([web](#), [pdf](#))
- [Slides](#) of the presentation at the *Rencontres R 2014* conference

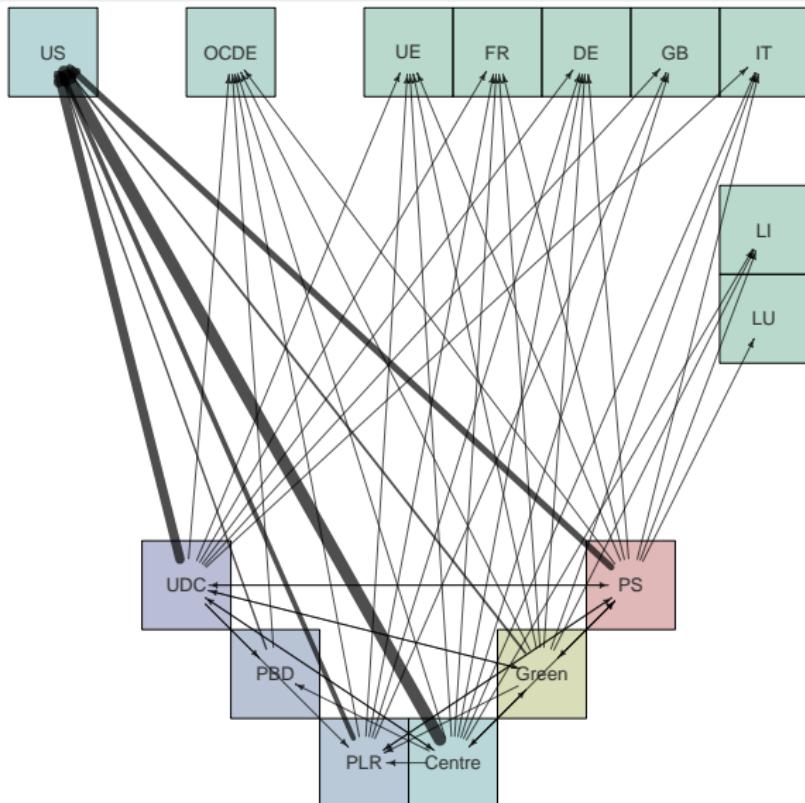
Development: The development version is hosted on the [R-Forge](#) platform.











Merci pour votre attention !

Contact: Marion.Deville@unige.ch

Package R 'spnet':

<http://emmanuel.rousseaux.me/r-package-spnet>

Package R 'debate':

<http://emmanuel.rousseaux.me/r-package-debate>

Understanding **surfers** through **twitter** (with **R**)

6th French R meeting | Anglet 2017
Yan Holtz, [@R_Graph_Gallery](#)



Understanding **surfers** through **twitter** (with **R**)

6th French R meeting | Anglet 2017
Yan Holtz, [@R_Graph_Gallery](#)





Developers



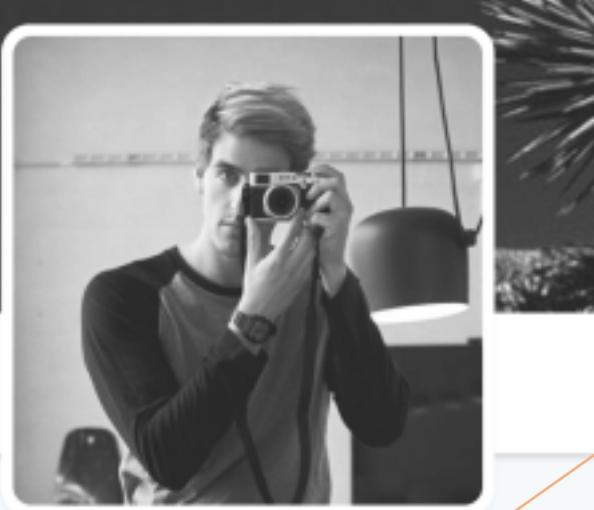
`library(twitteR)`

`searchTwitter("#surf")`

`for(i in seq(1,365)){`

Get additional
information

`Sys.sleep(86380)`



Yannick Wolff

@yannickwolffcom

Photographer | Instagram & Snapchat

@yannickwolff

📍 Darmstadt, Germany

```
"@yannickwolffcom$"getFollowers()
```

```
location("@yannickwolffcom")
```

```
library(ggmap)  
geocode("Darmstadt, Germany")
```



Yannick Wolff @yannickwolffcom · 16 mai

Beautiful dunes along Ireland's west coast.

#ireland #soirish #sogreen #dunes #bundoran #surf... instagram.com
[/p/BUJj5k2h4b3/](https://www.instagram.com/p/BUJj5k2h4b3/)

🌐 À l'origine en anglais

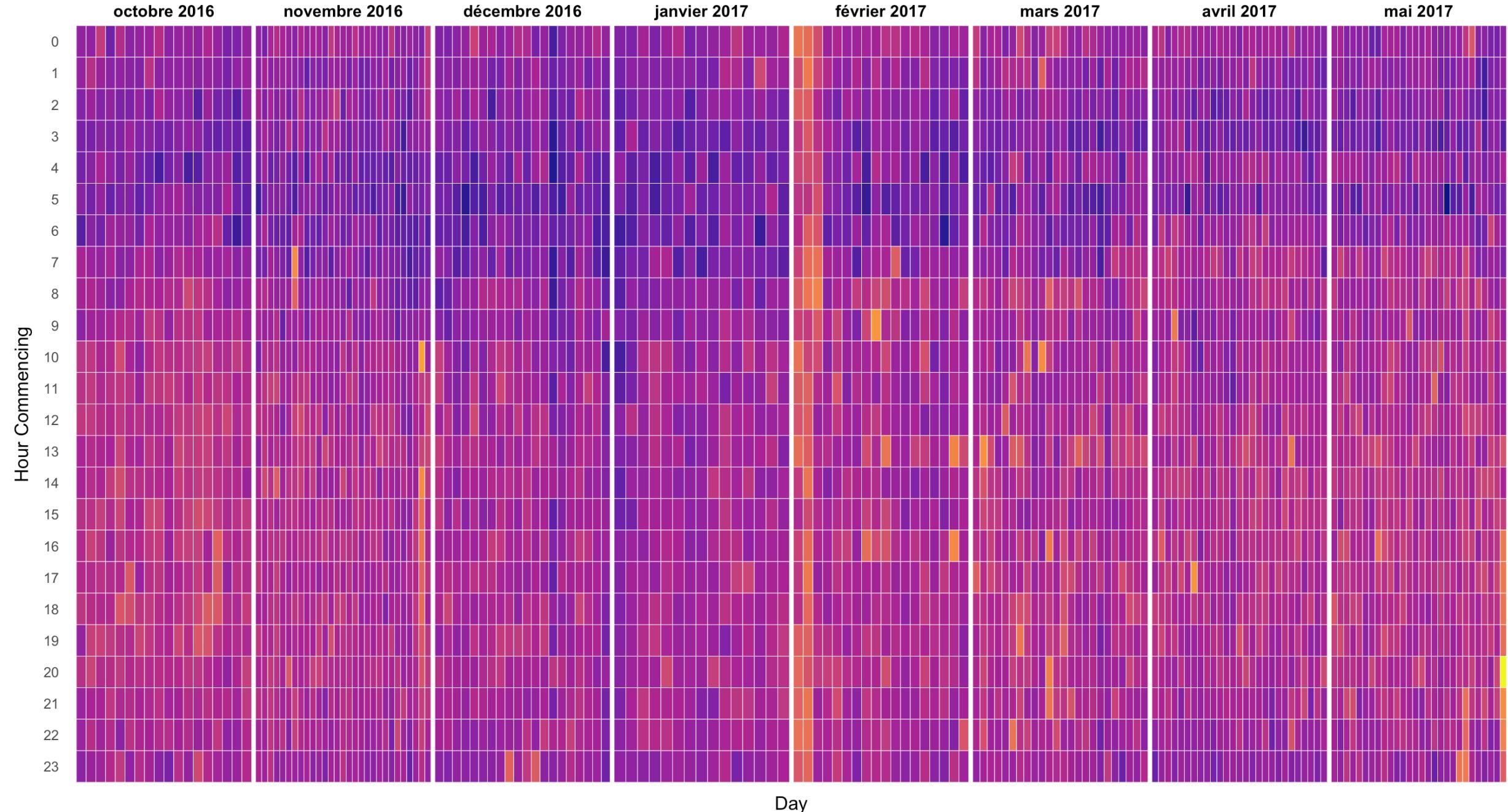


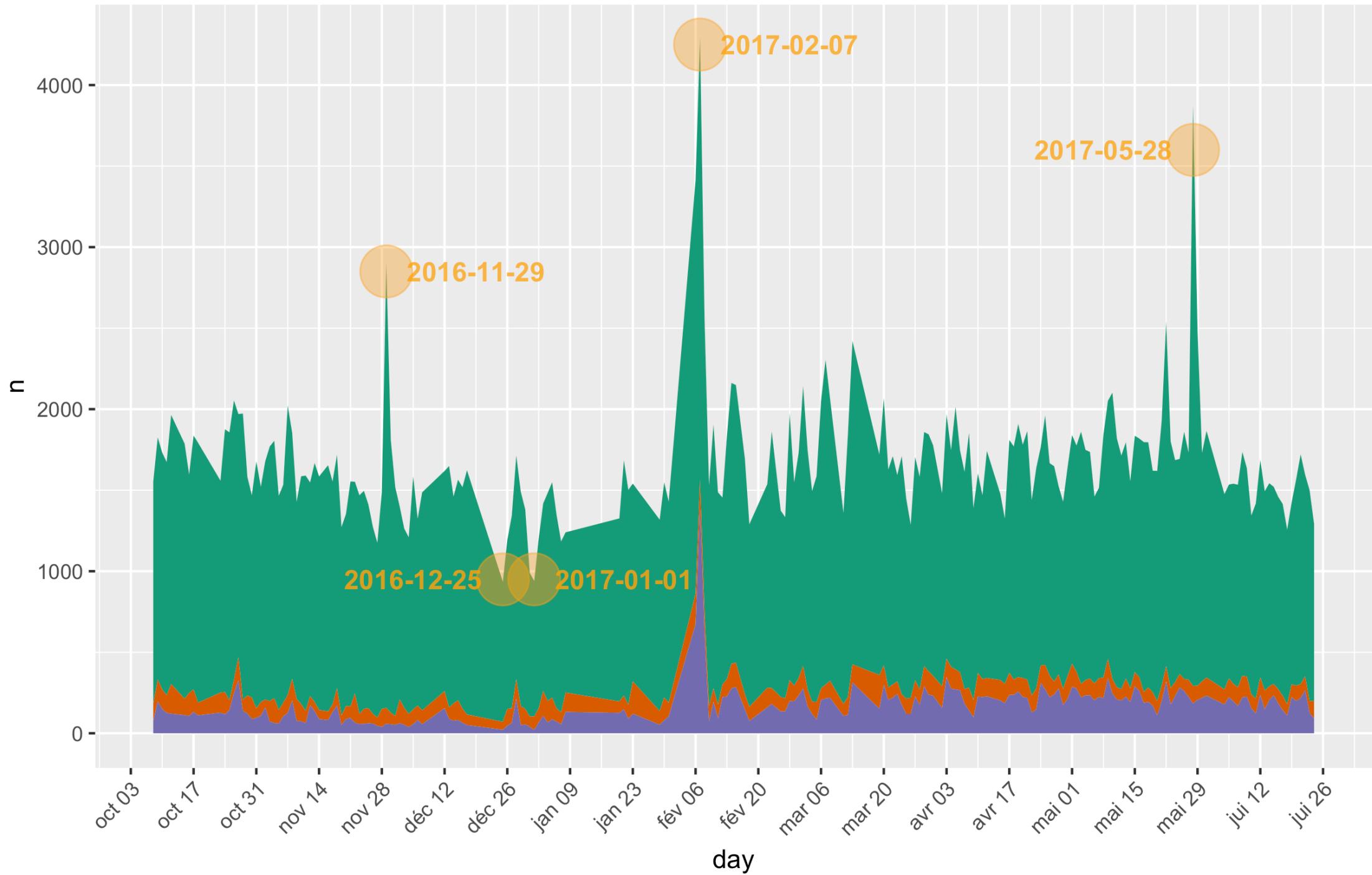
```
library(rworldmap)  
library(sp)  
SpatialPoints()
```

```
library(wordcloud2)
```

```
library(sp)  
spDistsN1(home, travel)
```

Hourly #surf tweet number





sport

- surf
- kite
- wind

nuestro beach derek
nuestra contina waves
ocean gmez arza
gmez medalla otro arza
sup video con plata
fem final masc vivas
sigue tgSURF 2016
reyes gana
teamesa indvzla

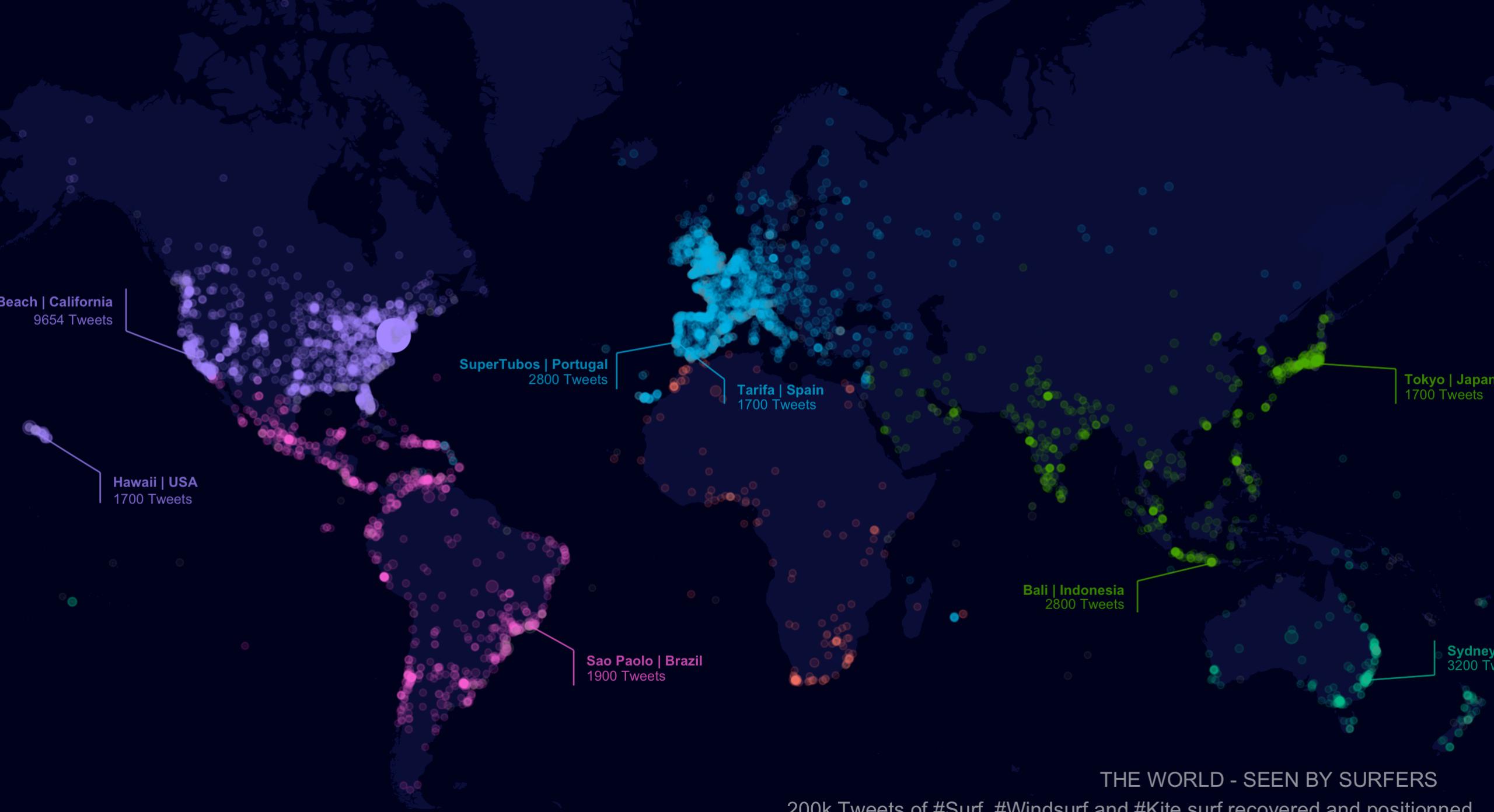


2016 shirt silver sonic 2014
plate northridge
board severne
mast size
records 1963
base 147 rock
adventure beach
amp live
outdoors
75m 101 surfs mens 2009
banzaipipeline north
summer cross



even today crazy
happy anniversary so chancedirect
way corzo best
time old
got ago
amp mundial
flies year jhony
realcottontale campen surfin confidence
years cre
thanks weve





THE WORLD - SEEN BY SURFERS
200k Tweets of #Surf, #Windsurf and #Kite surf recovered and positionned
www.r-graph-gallery.com

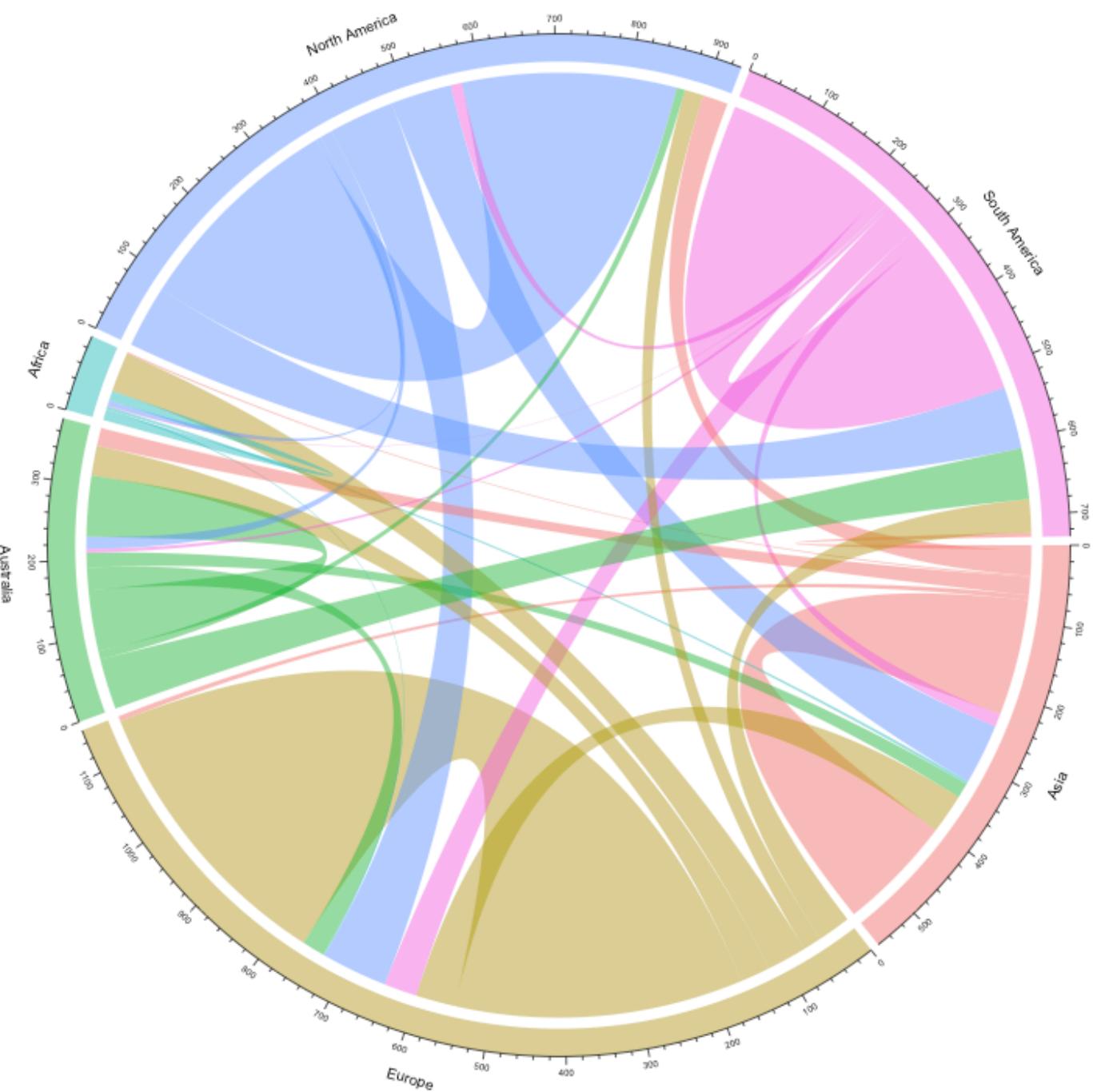
Where #SURFERS travel

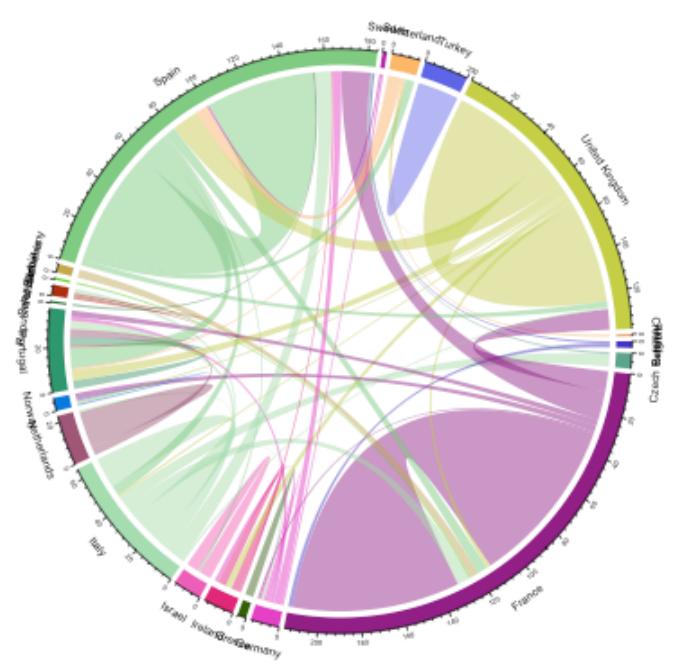
300k #Surf tweets recovered by

@R_Graph_Gallery

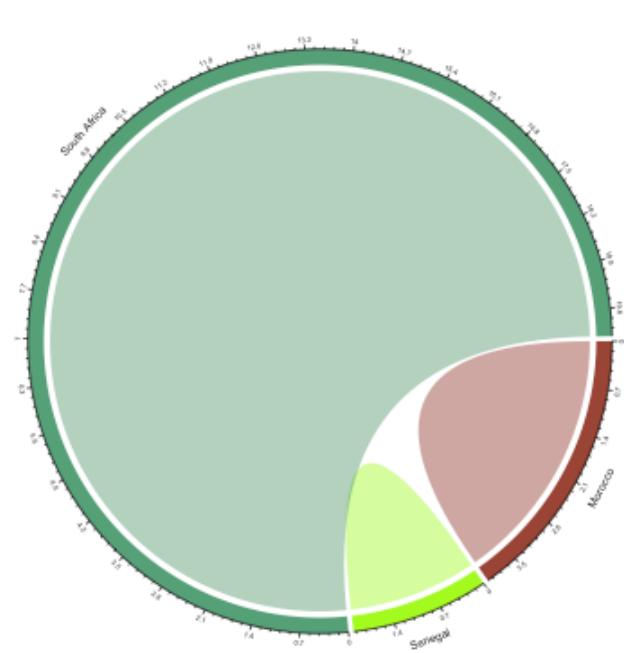
#Surfer flows between continents

300k tweets recovered by
[@R_Graph_Gallery](#) and visualized with R
and the circlize library.

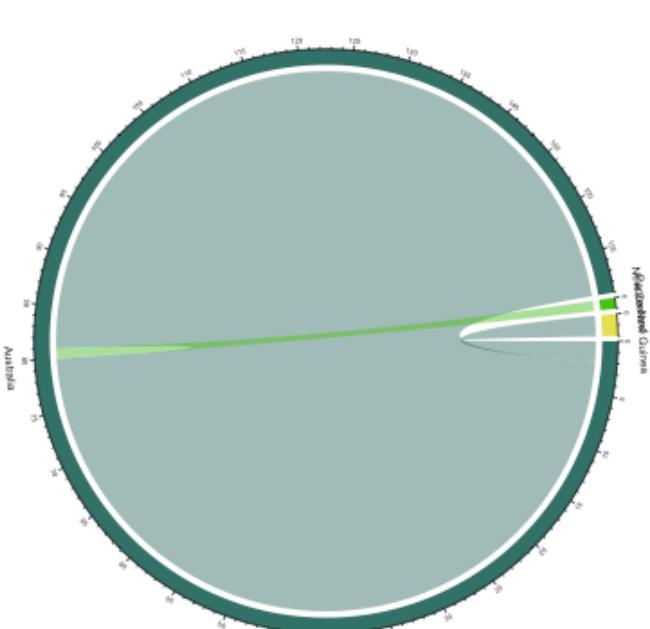




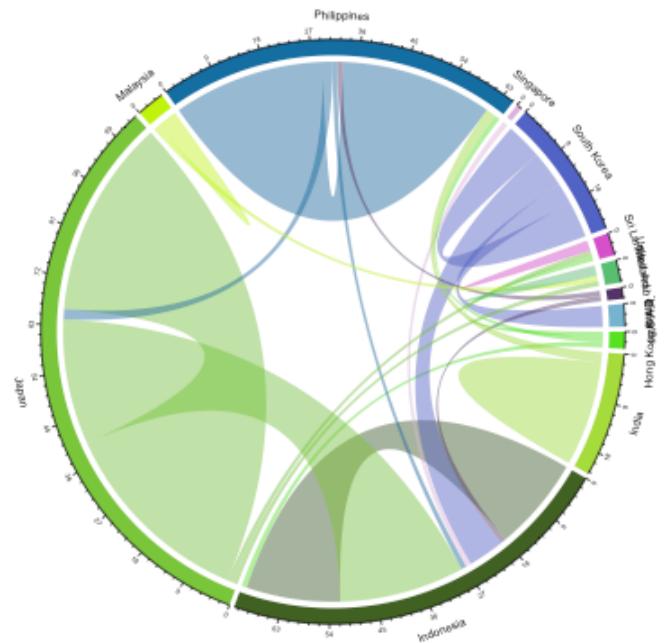
Europe



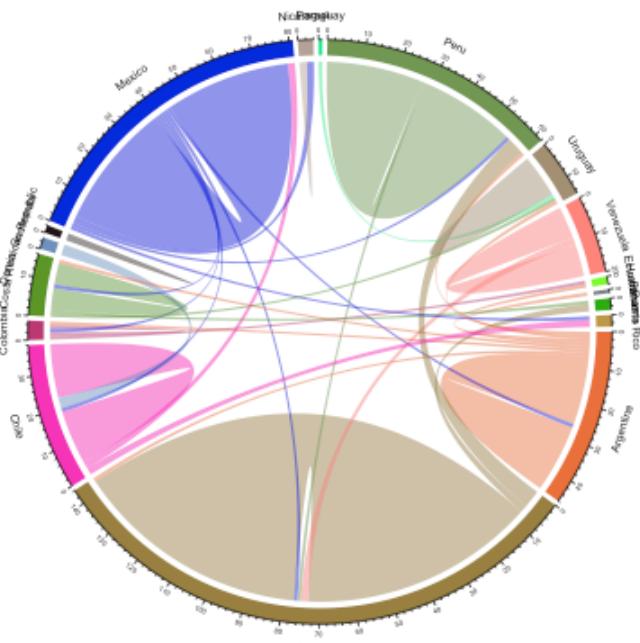
Africa



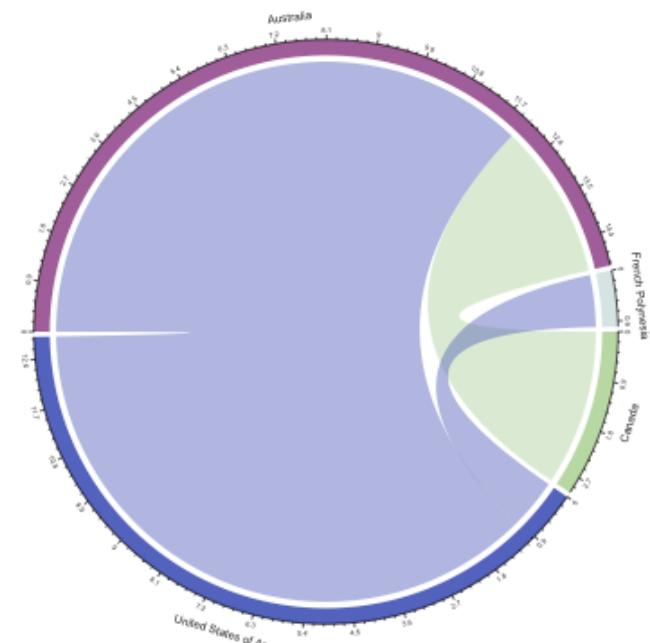
Antarctica



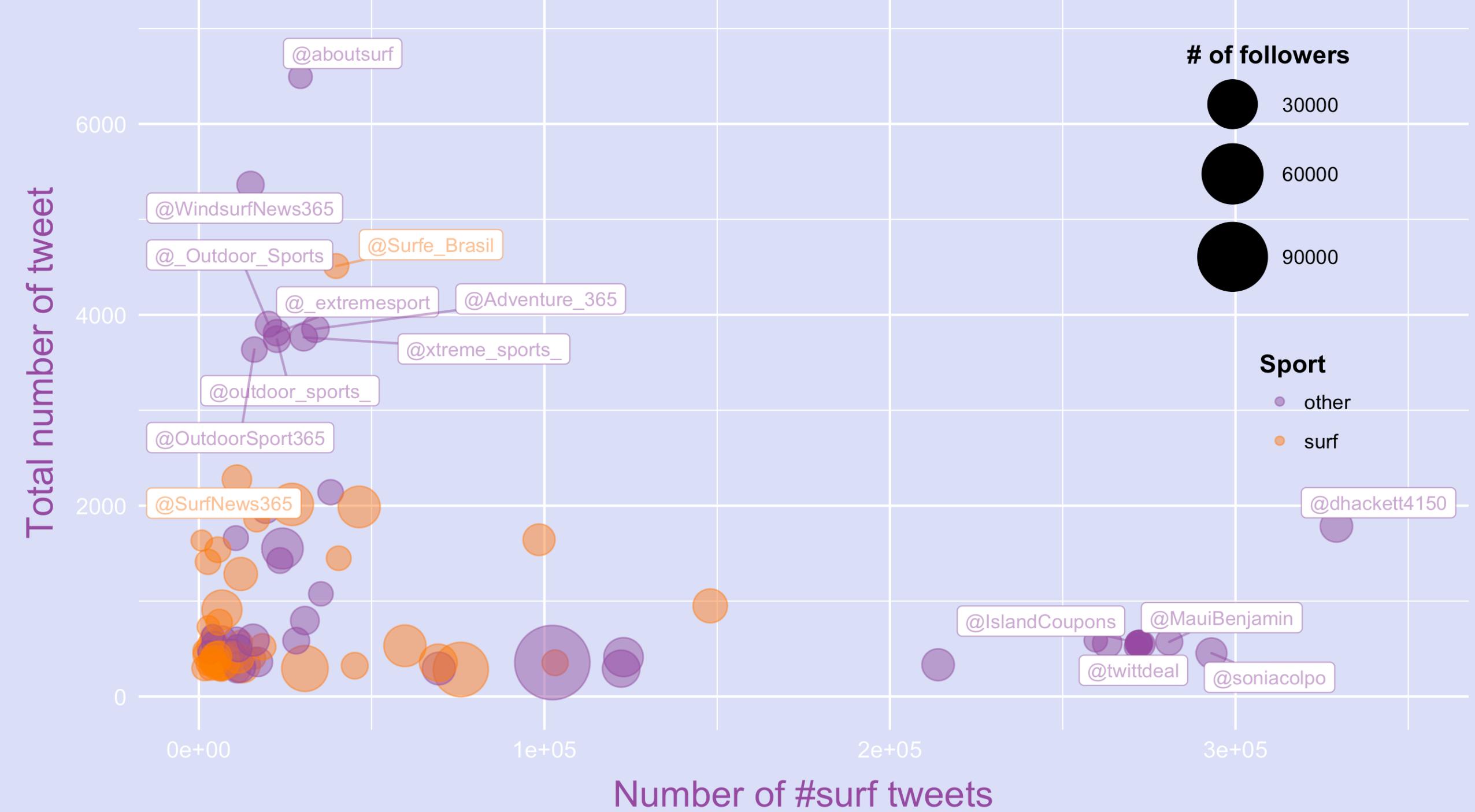
Asia

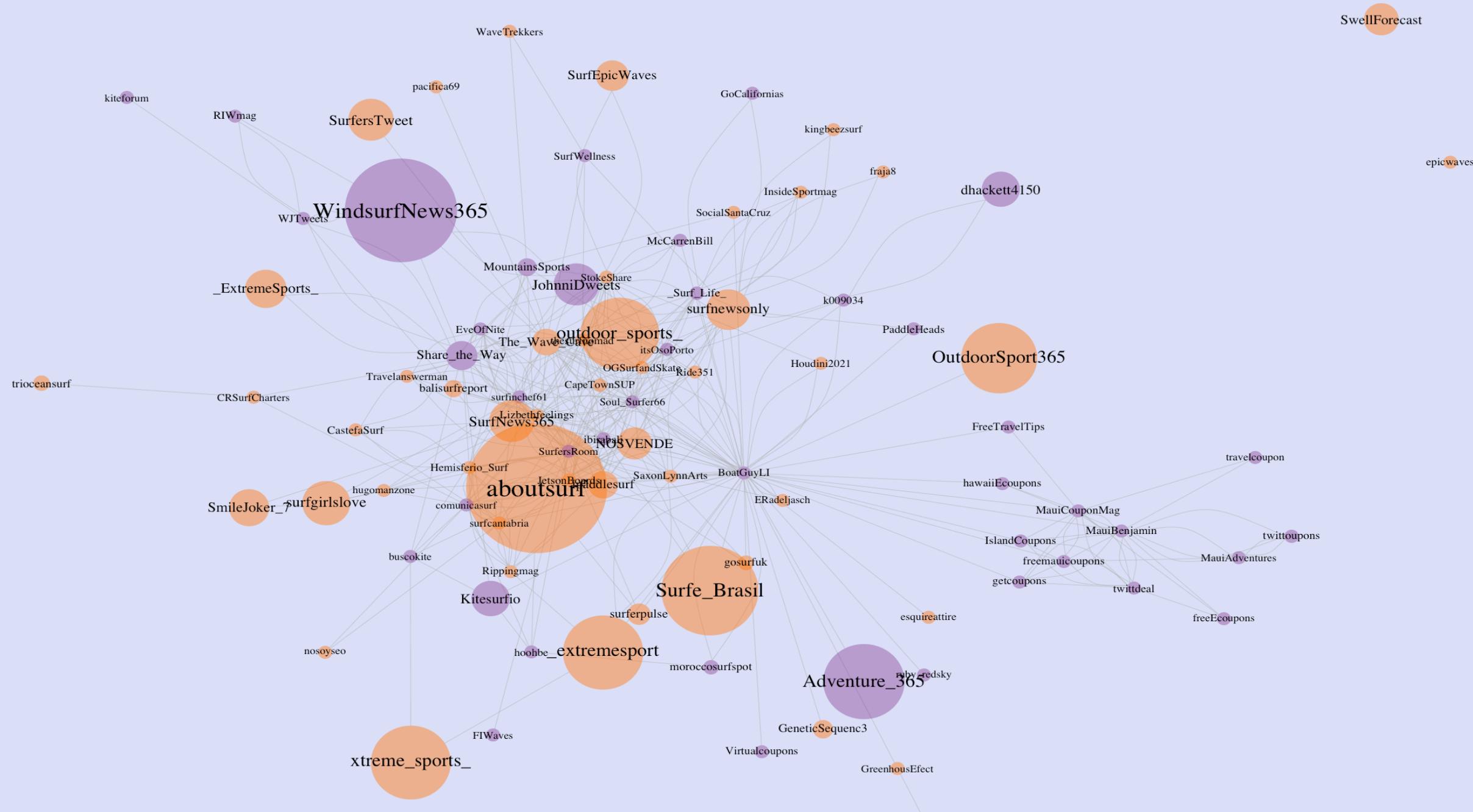


South America



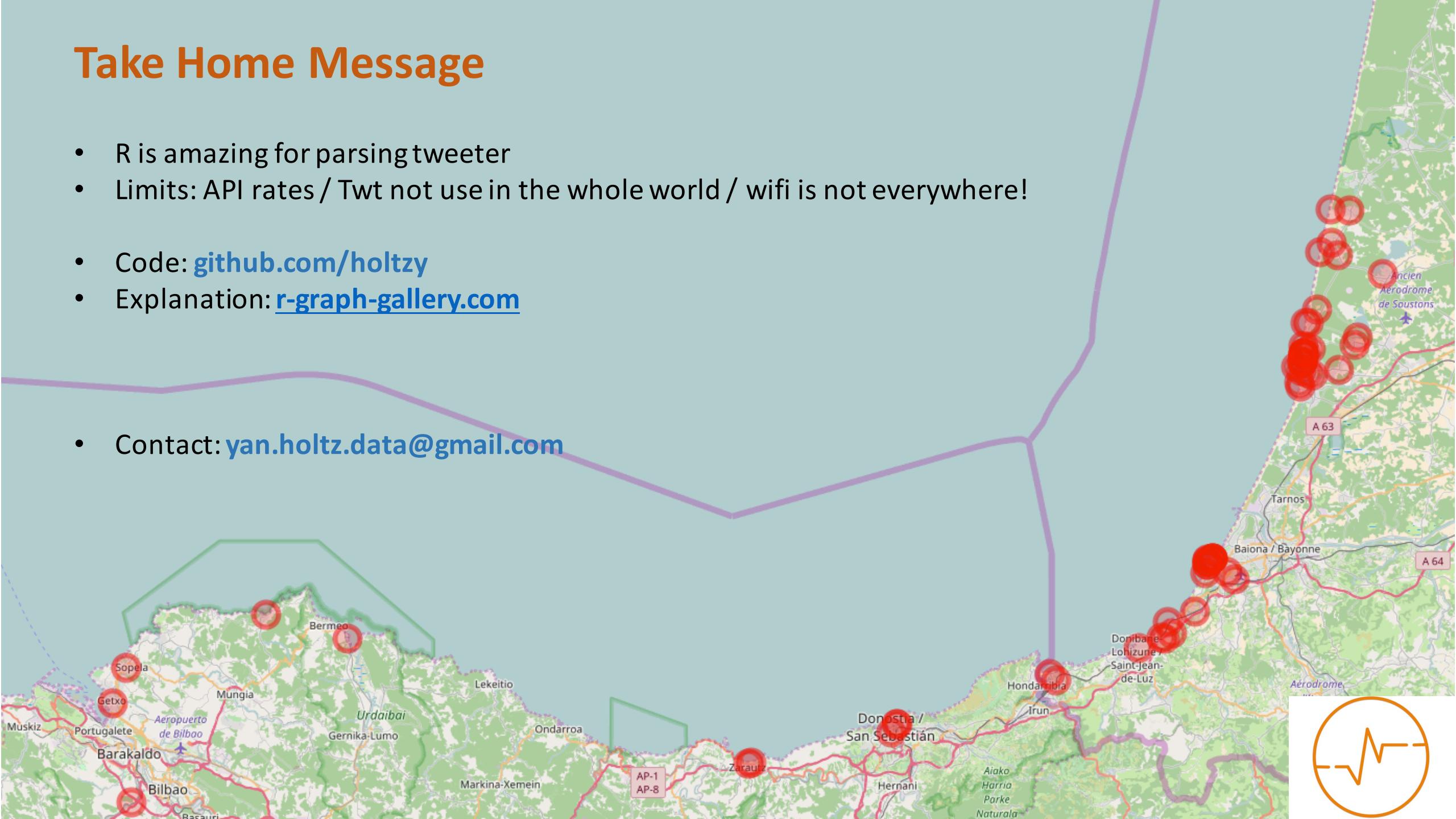
North America





Take Home Message

- R is amazing for parsing tweeter
- Limits: API rates / Twt not use in the whole world / wifi is not everywhere!
- Code: github.com/holtzy
- Explanation: r-graph-gallery.com
- Contact: yan.holtz.data@gmail.com



1 dataset, 10 dataviz

1 dataset, 10 dataviz

Pourquoi ?

How to Lie with Charts

By [Jack Dougherty](#), last updated March 25, 2017

One of the best ways to learn how to detect bias in data visualization is to intentionally manipulate a chart, and tell two (or more) opposing stories with the same data. You'll learn what to watch out for when viewing other people's charts, and think more carefully about the ethical issues when you design your own.

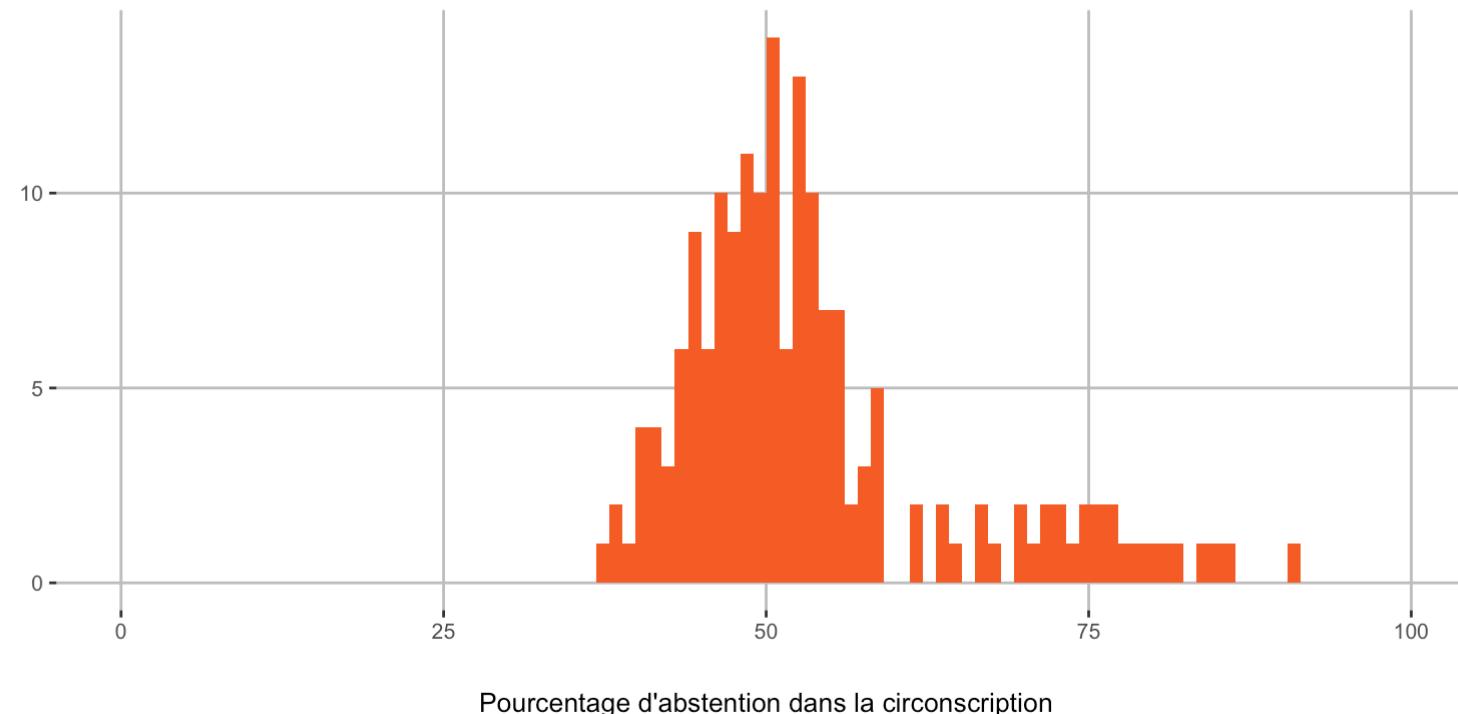
Le dataset

```
# devtools::install_github("ThinkRstat/legislatives2017")
legislatives2017::second_tour[1:5,]

## # A tibble: 5 x 21
##   code_dpt num_circ nom_dpt          nom_reg code_reg Inscrits
##   <chr>     <dbl>  <chr>          <chr>    <chr>     <dbl>
## 1 01         1     Ain Auvergne Rhone Alpes     84     82676
## 2 01         1     Ain Auvergne Rhone Alpes     84     82676
## 3 01         2     Ain Auvergne Rhone Alpes     84     93507
## 4 01         2     Ain Auvergne Rhone Alpes     84     93507
## 5 01         3     Ain Auvergne Rhone Alpes     84     75548
## # ... with 15 more variables: Abstentions <dbl>, Votants <dbl>,
## #   Blancs <dbl>, Nuls <dbl>, Exprimes <dbl>, candidat <chr>,
## #   Nuances <chr>, Voix <dbl>, p_inscrits <dbl>, p_exprimes <dbl>,
## #   resultat <chr>, civilité <chr>, Score <dbl>, p_abstentions <dbl>,
## #   ID <chr>
```

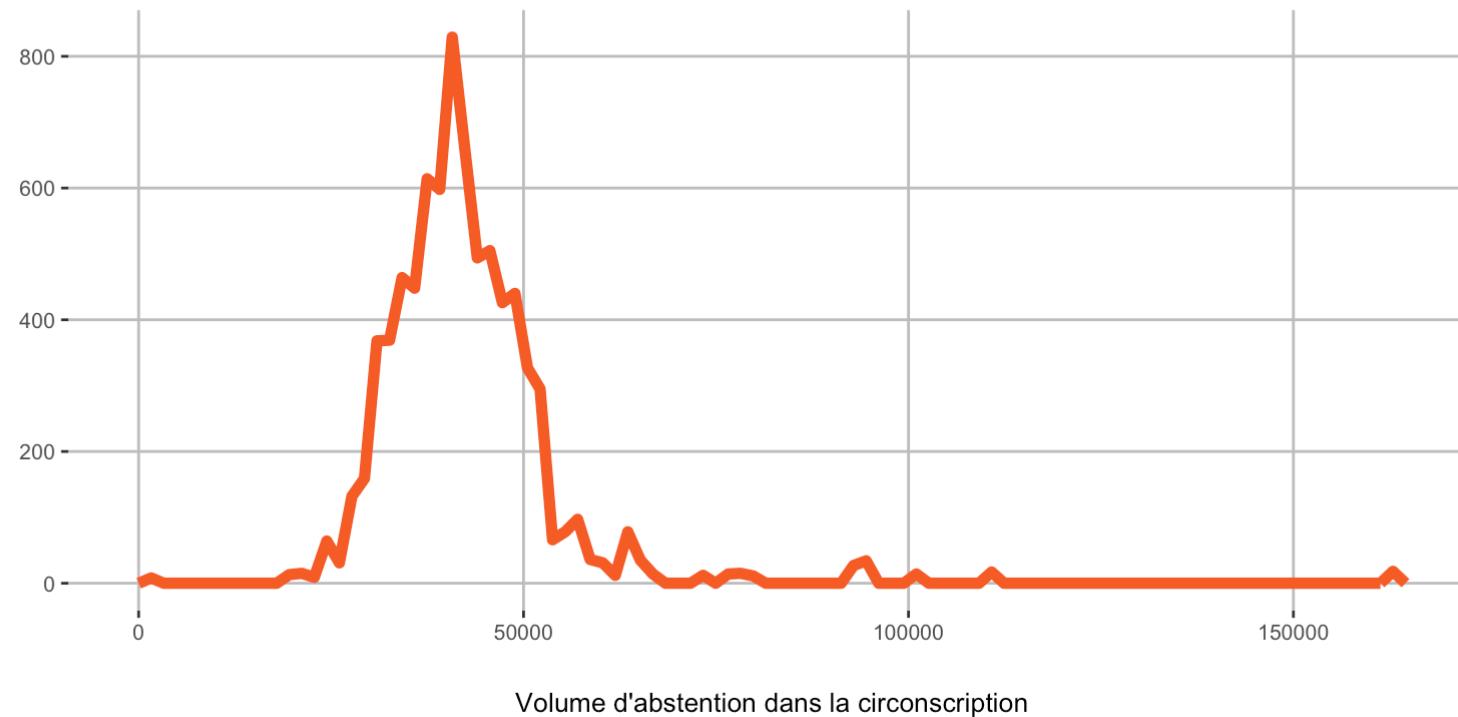
1. geom_hist()

L'abstention au premier tour des législatives 2017



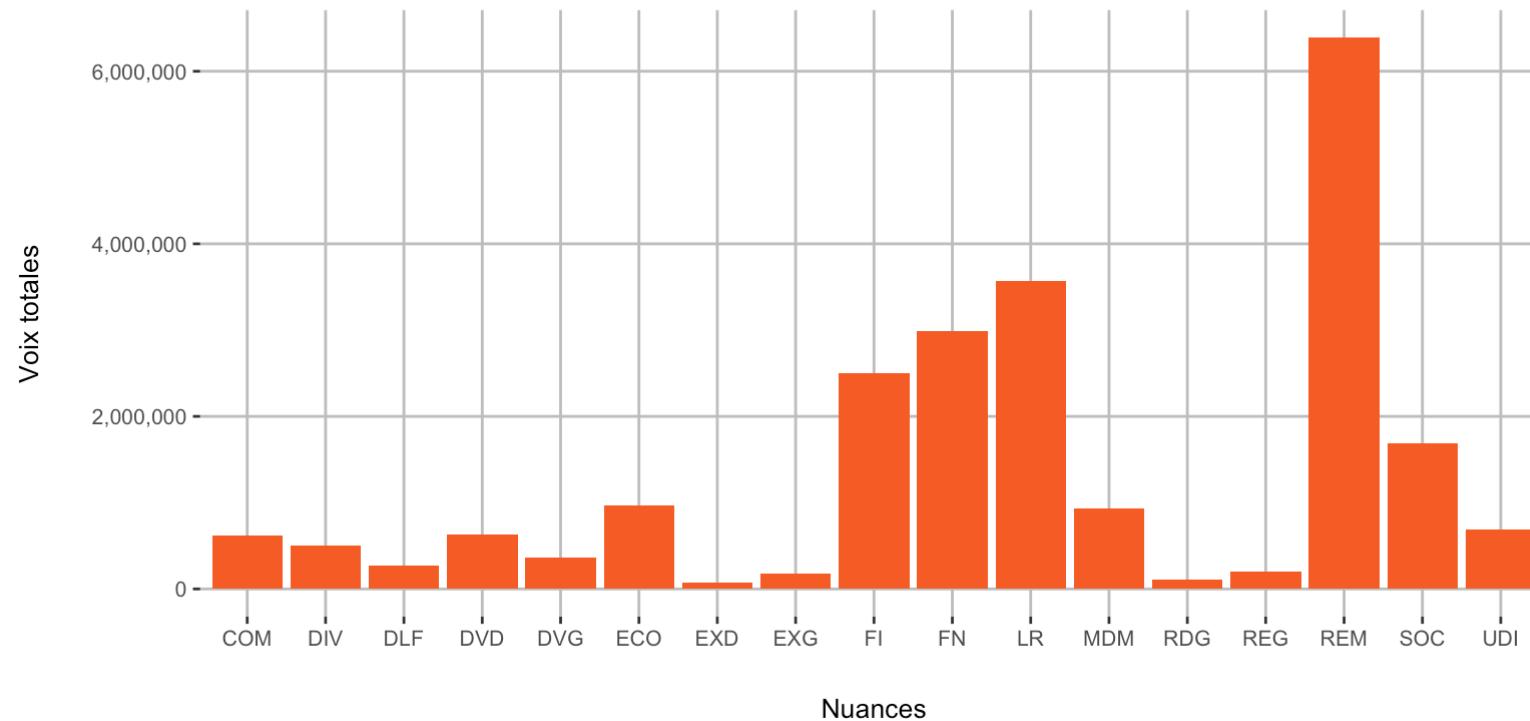
2. geom_freqpoly()

L'abstention au premier tour des législatives 2017



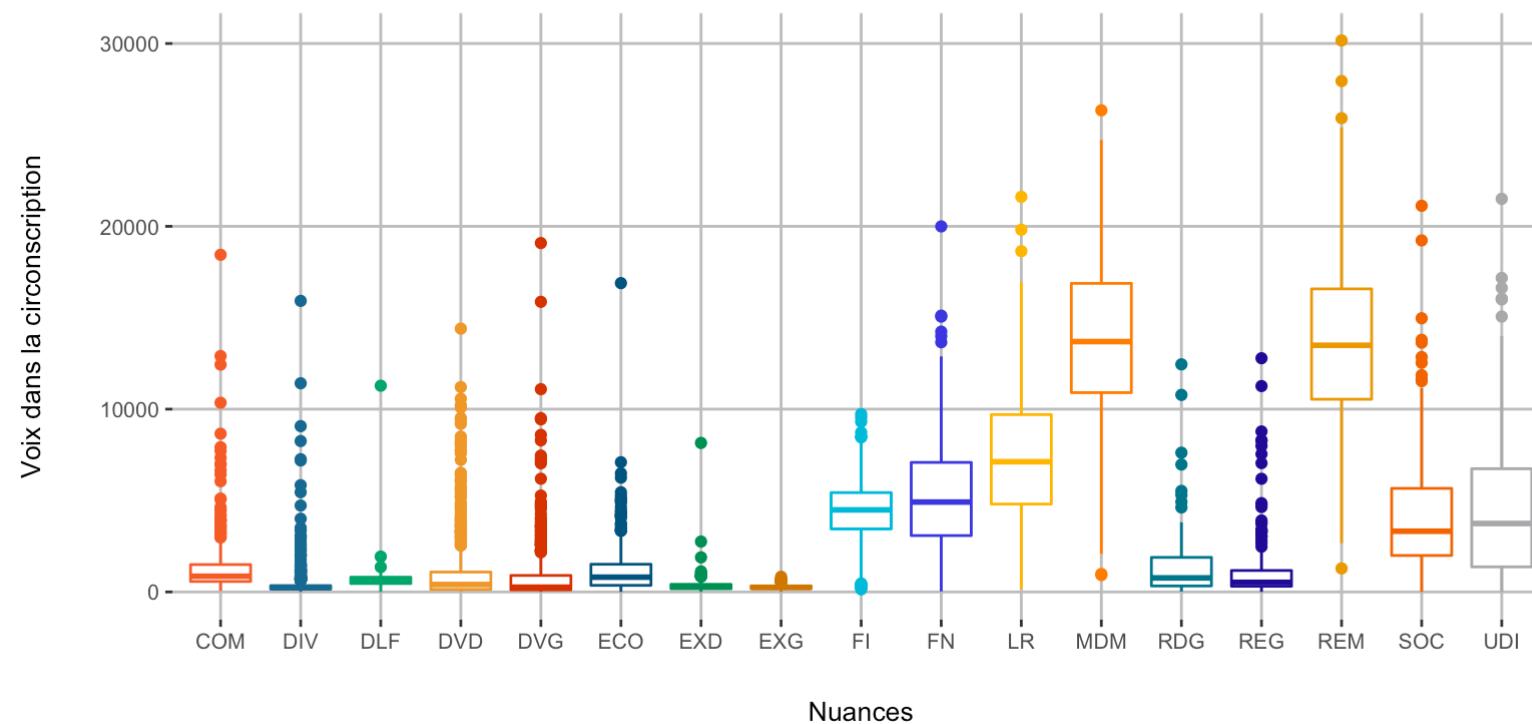
3. geom_bar()

Résultats du premier tour des législatives 2017



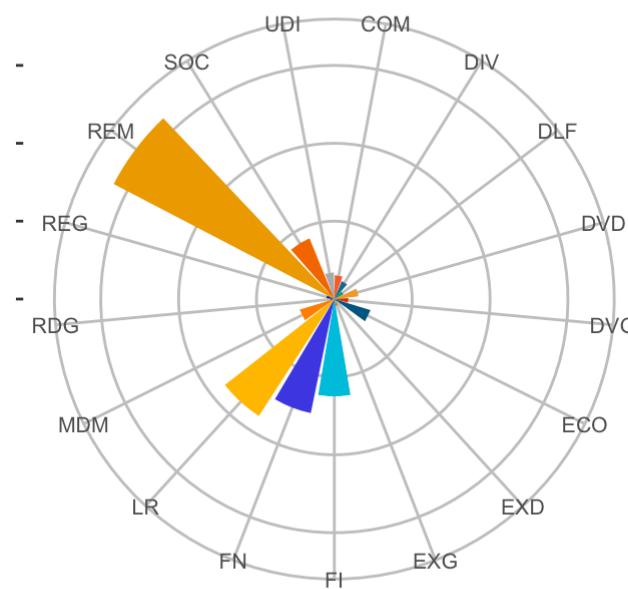
4. geom_boxplot()

Résultats du premier tour des législatives 2017

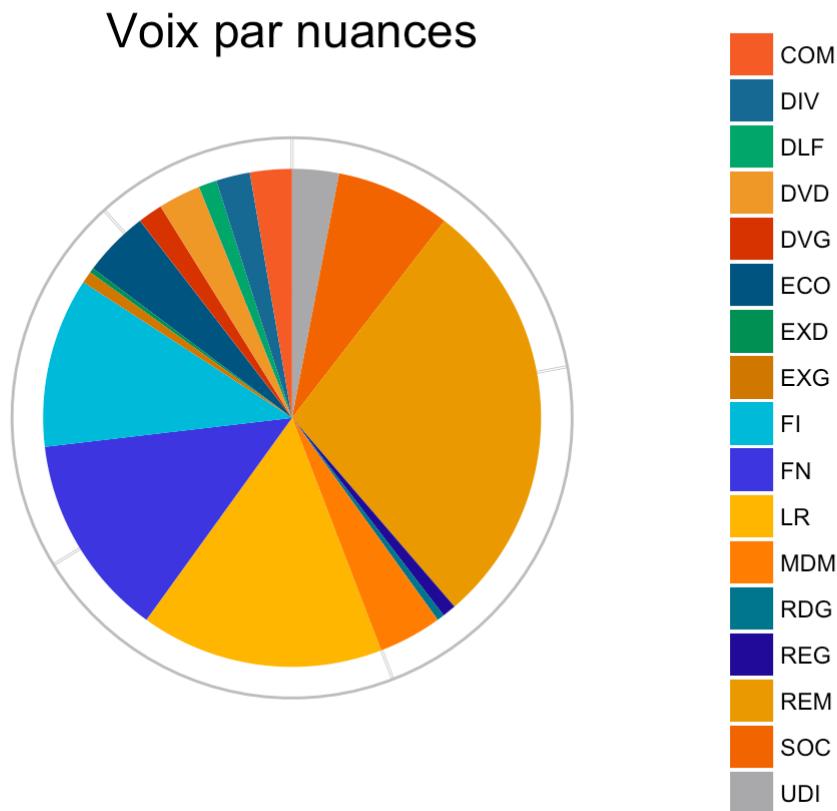


5. geom_bar() + coord_polar()

Voix par nuances au premier tour des législatives 2017

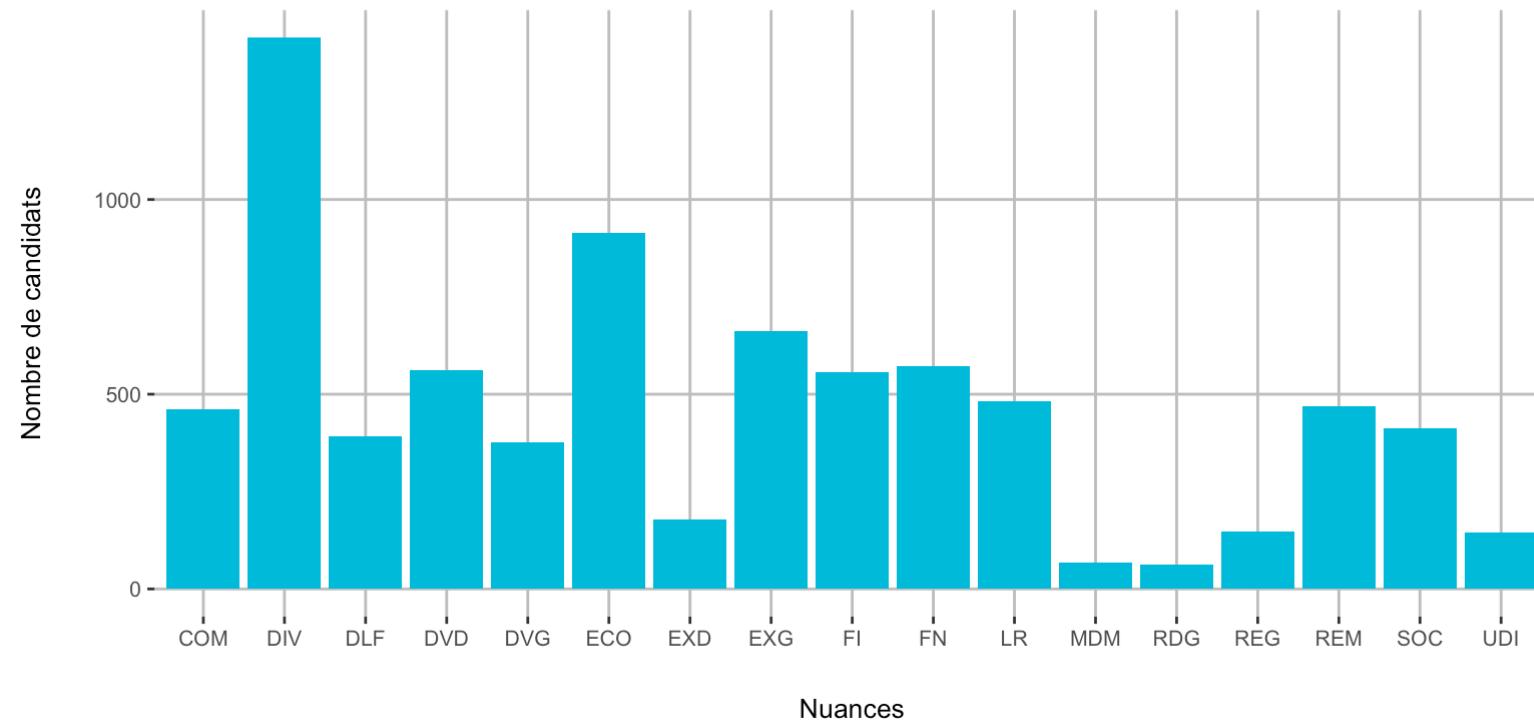


6. geom_bar() + coord_polar()



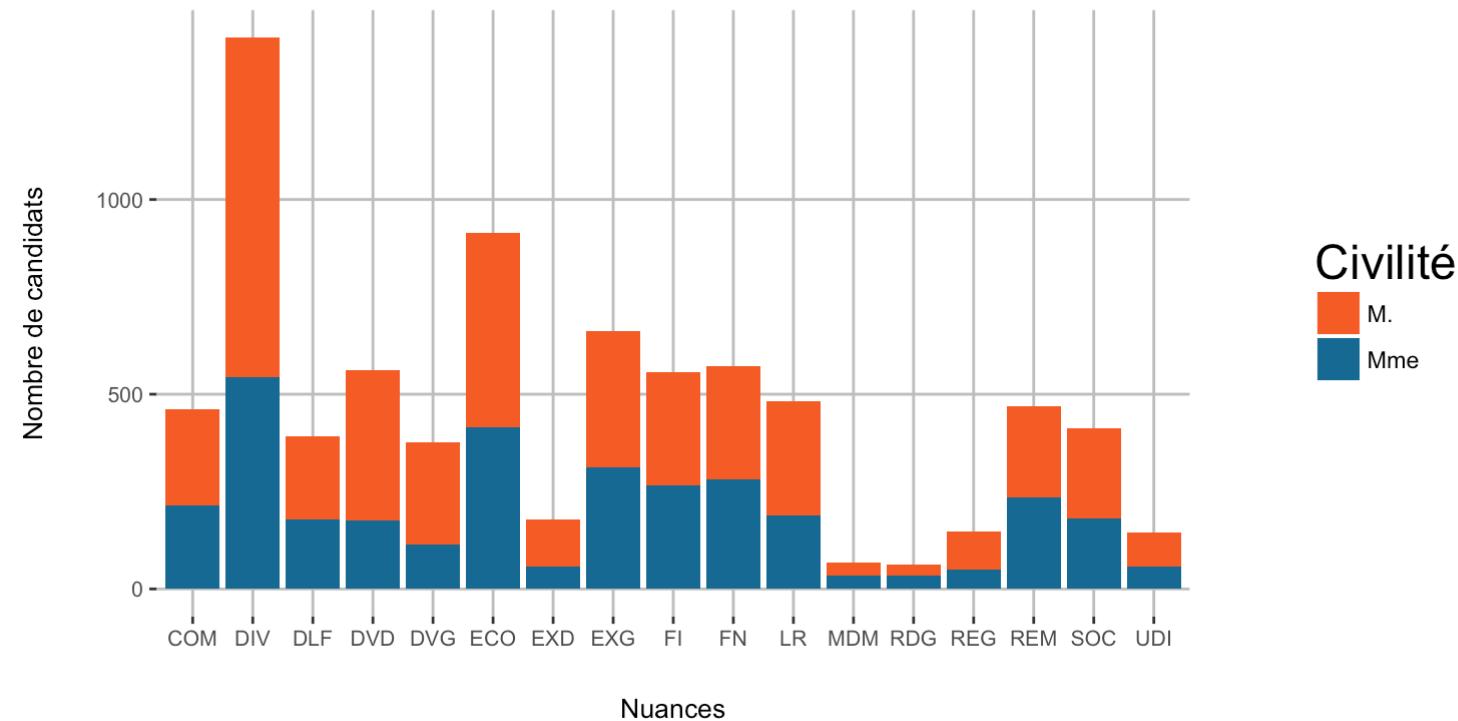
7. geom_bar()

Candidats par nuance au premier tour des législatives 2017

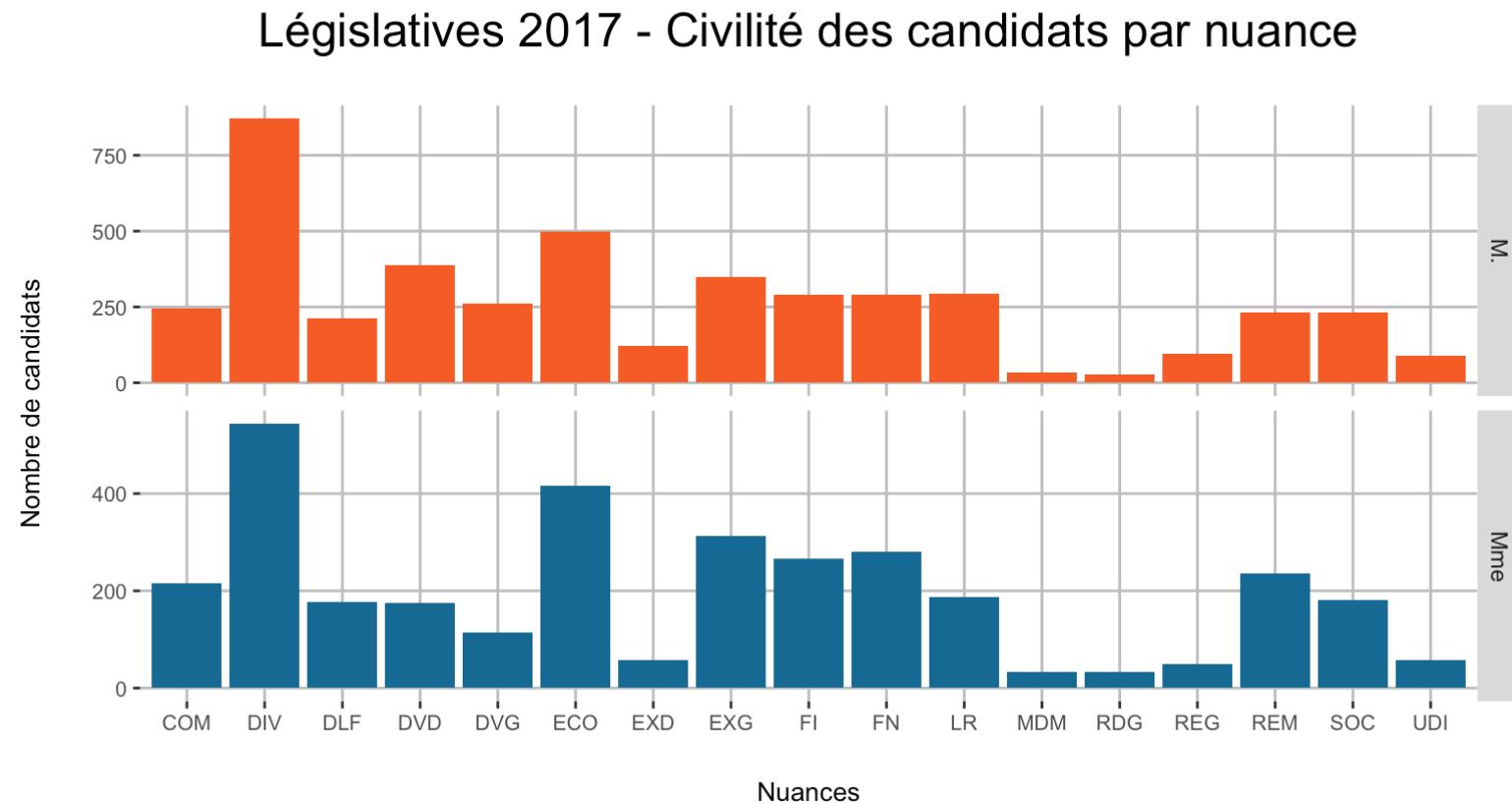


8. geom_bar(fill = civilite)

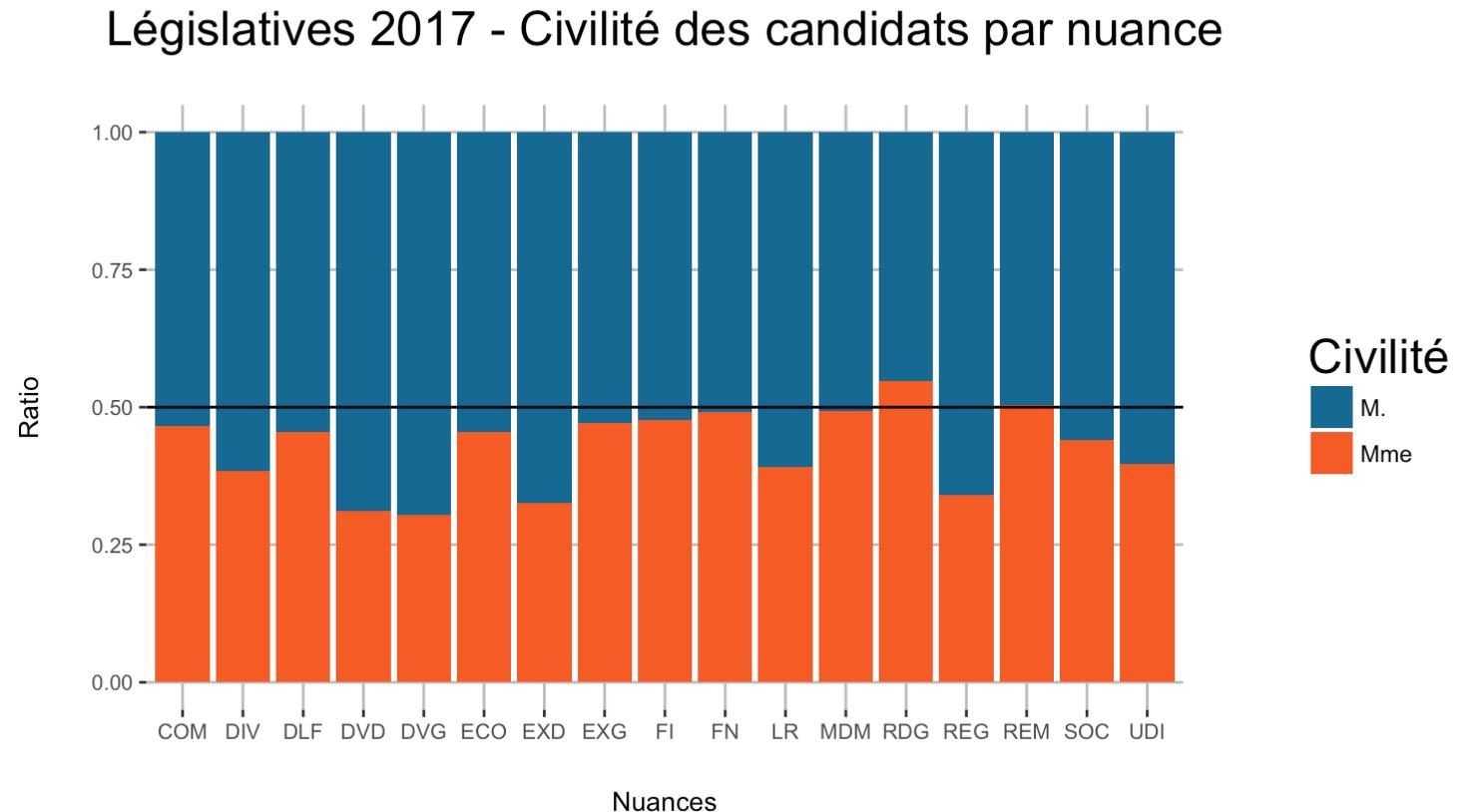
Candidats & civilité au premier tour des législatives, par nuances



9. geom_bar() + facet_wrap(scales = "free_y")



10. geom_bar('position = "fill")



Civilité

M.
Mme



Projet ART

« Apprentissage de R pour Tous »
« Aprendizaje de R para Todos »

Sixièmes Rencontres R, Anglet, juin 2017

Christian PAROISSIN
Maître de conférences HDR

Guillaume ISAAC
Ingénieur pédagogique

Juin 2017



Projet ART

« Apprentissage de R pour Tous »
« Aprendizaje de R para Todos »

Sixièmes Rencontres R, Anglet, juin 2017

Christian PAROISSIN
Maître de conférences HDR

Guillaume ISAAC
Ingénieur pédagogique

Juin 2017

Le contexte

Omniprésence des données → besoin de formations à R

- pour un public de spécialistes ou pas
- en formation initiale ou continue

Au démarrage : appel à projets internes UPPA en FOAD

- transformation d'un cours de M1 en présentiel en distanciel
- soutien du pôle ARTICE de la DN/UPPA
- plateforme Moodle

Pourquoi ne pas aller plus loin ?

Le contexte

Omniprésence des données → besoin de formations à R

- pour un public de spécialistes ou pas
- en formation initiale ou continue

Au démarrage : appel à projets internes UPPA en FOAD

- transformation d'un cours de M1 en présentiel en distanciel
- soutien du pôle ARTICE de la DN/UPPA
- plateforme Moodle

Pourquoi ne pas aller plus loin ?

Le dispositif pédagogique en ligne ART

Postulat de base : formation complémentaire pour l'apprenant

- concepts de statistique déjà connus
- mise en œuvre des méthodes statistiques avec R

Conséquence : un public très hétérogène... mais motivé !

Un dispositif modulaire : choix de 12 modules parmi 25

- deux types de modules : programmation ou analyse statistique
- sélection : par le responsable de la formation ou par l'apprenant

(Encore) un MOOC ? Non

Le dispositif pédagogique en ligne ART

Postulat de base : formation complémentaire pour l'apprenant

- concepts de statistique déjà connus
- mise en œuvre des méthodes statistiques avec R

Conséquence : un public très hétérogène... mais motivé !

Un dispositif modulaire : choix de 12 modules parmi 25

- deux types de modules : programmation ou analyse statistique
- sélection : par le responsable de la formation ou par l'apprenant

(Encore) un MOOC ? Non



Analyse pédagogique

Contenu d'un module

- note de cours (et plus tard : capsules vidéo)
- test formatif en ligne

Accès séquentiel aux ressources

- un module par semaine
- obligation d'une réussite supérieure à 75 % au test précédent

Trois séances en présentiel : début, milieu et fin

Evaluation : test sommatif en ligne



Analyse pédagogique

Contenu d'un module

- note de cours (et plus tard : capsules vidéo)
- test formatif en ligne

Accès séquentiel aux ressources

- un module par semaine
- obligation d'une réussite supérieure à 75 % au test précédent

Trois séances en présentiel : début, milieu et fin

Evaluation : test sommatif en ligne

Exploitations actuelles et futures

Dispositif actuellement utilisé par :

- étudiants de M1 MSID
- doctorants de l'ED SEA (biologie, chimie, physique, etc.)

Et très prochainement par :

- étudiants en géographie et en STAPS
- doctorants des universités transfrontalières de l'UPPA (Bilbao et Pampelune) : traduction partielle en espagnol

Mais encore par :

- personnel de l'UPPA (en lien avec la DRH)
- personnes extérieures à l'UPPA (en lien avec la FORCO)

Exploitations actuelles et futures

Dispositif actuellement utilisé par :

- étudiants de M1 MSID
- doctorants de l'ED SEA (biologie, chimie, physique, etc.)

Et très prochainement par :

- étudiants en géographie et en STAPS
- doctorants des universités transfrontalières de l'UPPA (Bilbao et Pampelune) : traduction partielle en espagnol

Mais encore par :

- personnel de l'UPPA (en lien avec la DRH)
- personnes extérieures à l'UPPA (en lien avec la FORCO)

En conclusion : ART, en trois mots...

1. **Modularité** : une formation à la carte, selon les connaissances et les besoins de l'apprenant
2. **Transdisciplinarité** : un public varié et motivé par un même objectif, l'acquisition de compétences en R
3. **Transfrontalier** : déploiement dans les universités transfrontalières et partenaires de l'UPPA

En conclusion : ART, en trois mots...

1. **Modularité** : une formation à la carte, selon les connaissances et les besoins de l'apprenant
2. **Transdisciplinarité** : un public varié et motivé par un même objectif, l'acquisition de compétences en R
3. **Transfrontalier** : déploiement dans les universités transfrontalières et partenaires de l'UPPA

Christian PAROISSIN
cparoiss@univ-pau.fr

<http://cparoiss.perso.univ-pau.fr/>



Christian PAROISSIN
cparoiss@univ-pau.fr

<http://cparoiss.perso.univ-pau.fr/>



Efficient simulation of complex queueing systems with the R package queuecomputer

Anthony Ebert, Paul Wu, Kerrie Mengersen, Fabrizio Ruggeri

2017/06/29

Efficient simulation of complex queueing systems with the R package queuecomputer

Anthony Ebert, Paul Wu, Kerrie Mengersen, Fabrizio Ruggeri

2017/06/29

Simulation efficace des files d'attente complexe avec le paquet R “queuecomputer”



ACEMS

AUSTRALIAN RESEARCH COUNCIL CENTRE OF EXCELLENCE FOR
MATHEMATICAL AND STATISTICAL FRONTIERS



La file d'attente à une boulangerie



Simple example (input)

```
library(queuecomputer)
library(dplyr)

n_customers <- 40
arrivals <- rexp(n_customers, 3.1) %>% cumsum()
service <- rexp(n_customers, 1)

n_servers <- 3
departures <- queue_step(arrivals, service, n_servers)

#n_servers <- as.server.stepfun(10, c(2,4))
```

Simple example (output)

```
head(arrivals, 3)
```

```
## [1] 0.2436070 0.6247821 0.6717843
```

```
head(service, 3)
```

```
## [1] 1.079881 1.028247 1.292262
```

```
head(departures$departures, 3)
```

```
## [1] 1.323488 1.653029 1.964046
```

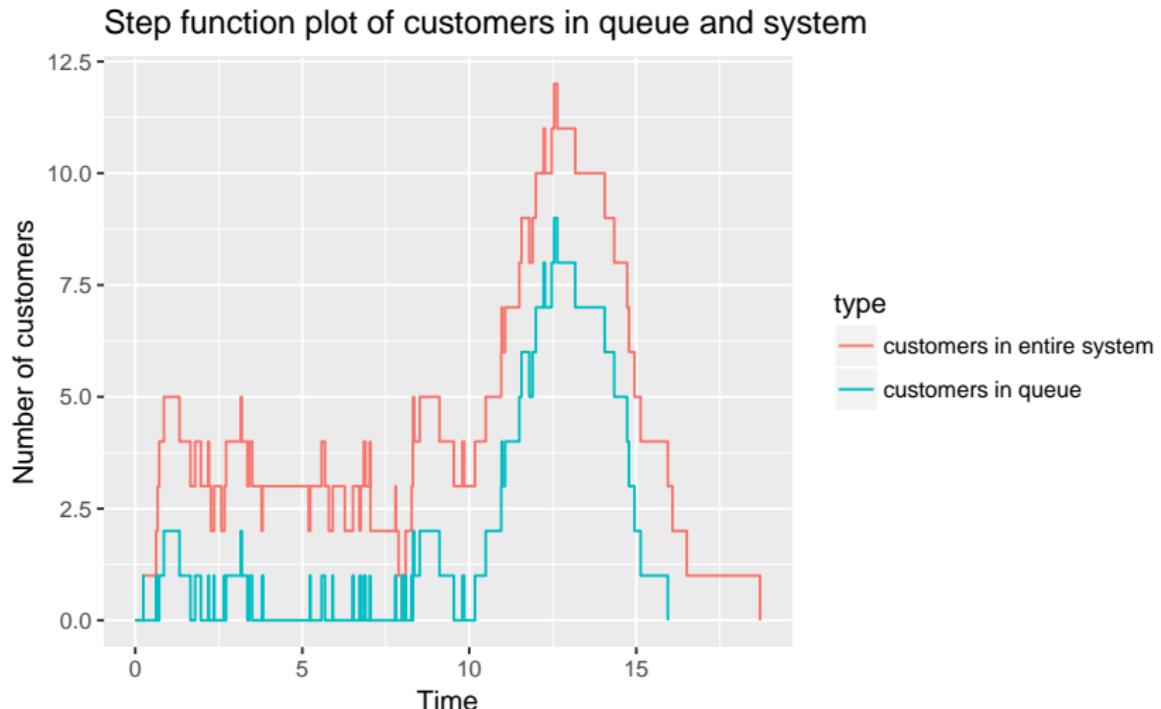
Summary

```
summary(departures)
```

```
## Total customers:  
## 40  
## Missed customers:  
## 0  
## Mean waiting time:  
## 0.783  
## Mean response time:  
## 1.98  
## Utilization factor:  
## 0.85  
## Mean queue length:  
## 1.96  
## Mean number of customers in system:  
## 4.22
```

Plot of number of customers in queue

```
plot(departures, which = 4)
```



Plot of customer status

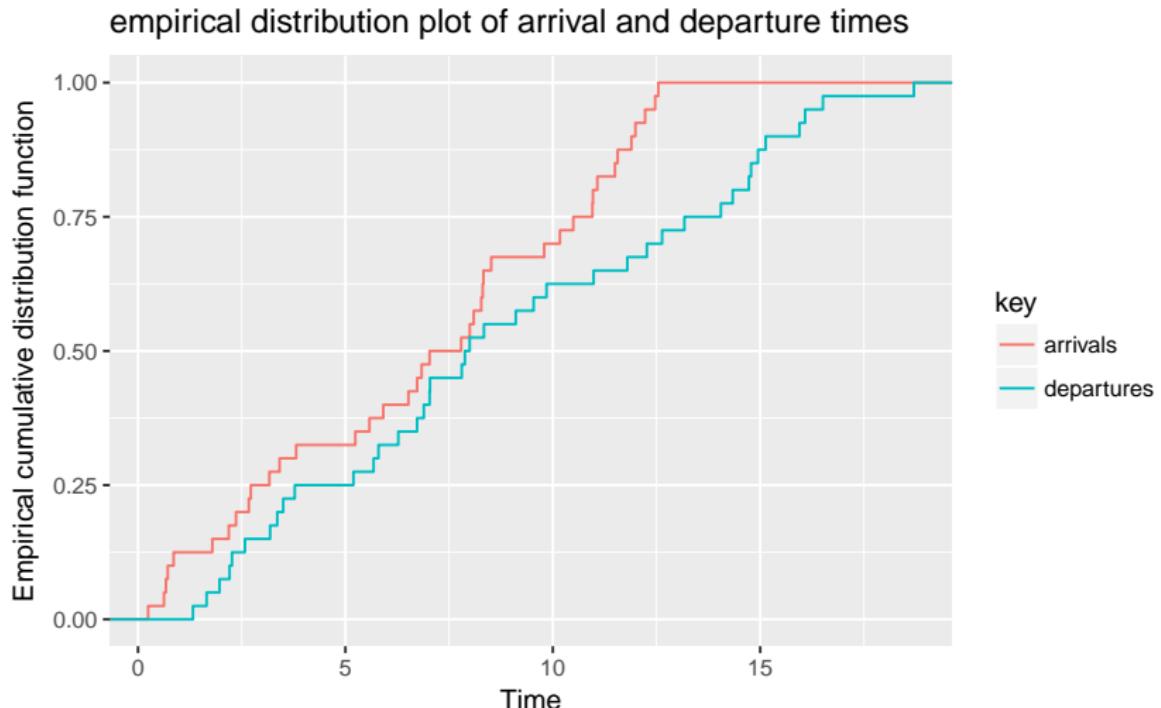
```
plot(departures, which = 5)
```

Line range plot of customer and server status

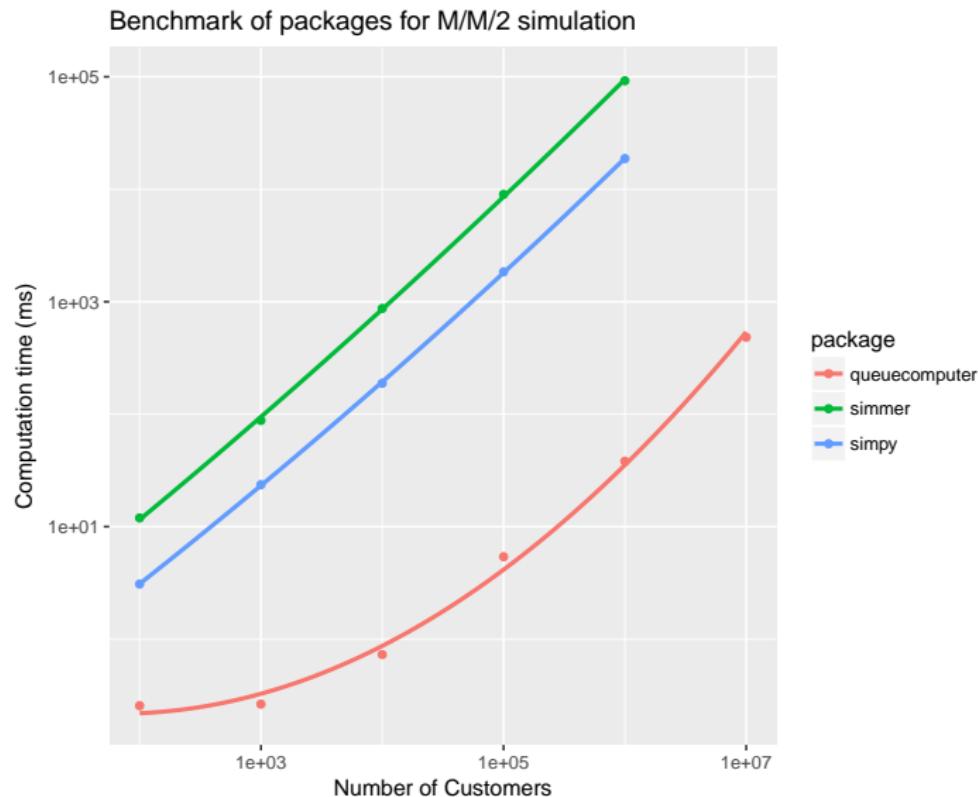


Empirical distribution of arrival and departure times

```
plot(departures, which = 6)
```



Temps de calcul / Computation time



Airport terminal

```
R> Passenger_df  
  
## # A tibble: 25,012 × 7  
##       ID FlightNo arrival route_imm arrive_imm service_imm  
##   <chr>    <fctr>   <dbl>    <fctr>      <dbl>      <dbl>  
## 1 Cordell, Megan ABI481  564.85 manual  566.8549  0.29075606  
## 2 Matheson, Dylan ABI481  564.85 manual  566.8532  0.15927226  
## 3 Avitia, Renee ABI481  564.85 manual  567.2014  0.22450319  
## 4 Woods, Tyrel  ABI481  564.85 smart gate 566.8377  0.18222445  
## 5 Pope, Christiana ABI481  564.85 smart gate 566.0994  0.09031344  
## 6 Espinoza, Mariah ABI481  564.85 smart gate 566.8928  0.43900281  
## 7 Pacheco, Charleen ABI481  564.85 manual  567.5558  0.12917143  
## 8 Harmon, Brendan ABI481  564.85 smart gate 566.3114  0.30565961  
## 9 William, Gerardo ABI481  564.85 smart gate 567.2563  0.31975687  
## 10 Hood, Colen  ABI481  564.85 smart gate 567.2181  0.33944458  
## # ... with 25,002 more rows, and 1 more variables: bag_time <dbl>
```

Airport terminal

```
server_df
```

```
## # A tibble: 2 × 2
##   route_imm    servers
##       <chr>     <list>
## 1 smart gate  <dbl [1]>
## 2 manual      <S3: list>
```

```
Passenger_df %>%
  left_join(server_df) %>%
  group_by(route_imm) %>%
  mutate(
    departures = queue(
      arrive_imm, service_imm, servers[[1]]))
)
```

Merci!

```
finish_presentation()
```

```
## [1] "Merci!"
```