

### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Mise à niveau R

Eric Marcon

10 November 2021



Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

## Installation



# R doit être installé proprement

Fric Marcon

Installation Démarrage

Rester à jour : version mineure en cours, ex.: 4.1

version\$version.string

## [1] "R version 4.1.2 (2021-11-01)"

Mettre à jour RStudio à chaque mise à jour de R



Référence : Travailler avec R, chapitre 1

Données

Fonctions

Structures de contrôle

Graphiques de

Tidyverse Graphiques

avec ggplot Aller plus loin



### Choisir un dossier de travail

### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Le dossier Home : ~



Toujours travailler dans un projet : File / New Project...

Ne jamais utiliser setwd() pour définir le dossier de travail: c'est toujours celui du projet.



# **Packages**

### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Les packages étendent les possibilités de R.

Installation depuis un dépôt officiel avec contrôle de qualité : CRAN.

Menu Tools / Install Packages...

Les packages sont installés dans ~/R/win-library/4.1/



Attention aux sauvegardes automatiques



Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

# Démarrage



### Les fenêtres de RStudio

#### Eric Marcon

Installation

### Démarrage

Données

Fonctions

Structures de contrôle

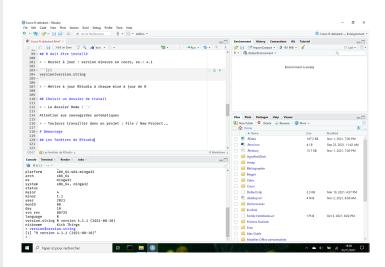
Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin





## Les variables

### Eric Marcon

Installation

### Démarrage

Données

Fonctions

Structures de

contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

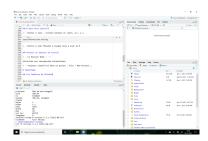
Aller plus loin

### Ecrire en haut à gauche.

```
# Affectation
a <- 1
# ou encore
1 -> a
# mais éviter
a = 1
```

Exécuter (Ctrl+Entrée) : voir en bas à gauche.

Environnement en haut à droite.





#### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

# Données



# R manipule des vecteurs

#### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Plutôt que des nombres

```
x <- 1:5
2 * x
```

## [1] 2 4 6 8 10

sqrt(x)

## [1] 1.000000 1.414214 1.732051 2.000000 2.236068



# Créer des vecteurs (1)

### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Des valeurs :

(x < -1)

## [1] 1

(x <- c("pommes", "poires"))</pre>

## [1] "pommes" "poires"

(x <- c(TRUE, FALSE))

## [1] TRUE FALSE



# Créer des vecteurs (2)

#### Eric Marcon

Installation

### Démarrage

## Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Une séquence :

(x <- 1:5)

## [1] 1 2 3 4 5

 $(x \leftarrow seq(from = 1, to = 5, by = 1))$ 

## [1] 1 2 3 4 5



# Créer des vecteurs (3)

#### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle

Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

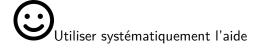
### Une répétition :

```
(x \leftarrow rep(1, 5))
```

## [1] 1 1 1 1 1

$$(x \leftarrow rep(1:2, each = 2))$$

## [1] 1 1 2 2



?rep

# Sélectionner des éléments (1)

#### Eric Marcon

Installation

Démarrage

#### Données

Fonctions

Structures de contrôle
Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

### Utiliser les crochets :

```
x <- (1:10) * 10
x[3]
```

## [1] 30

x[-5]

**##** [1] 10 20 30 40 60 70 80 90 100

Utiliser des vecteurs pour sélectionner :

x[c(1, 3)]

## [1] 10 30



## Tester des éléments

### Eric Marcon

Installation Démarrage

### Données

Fonctions

Structures de

contrôle Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

Tirer des nombres dans une séquence, trouver lesquels sont pairs.

```
x <- 1:100
# Echantillonnage
(y <- sample(x, 5))
## [1] 91 16 67 14 18</pre>
```

```
## [1] FALSE TRUE FALSE TRUE TRUE
```

(y%2 == 0)



# Sélectionner des éléments (2)

#### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de

contrôle Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

### Utiliser les crochets :

y[y<mark>\%2 == 0</mark>]

## [1] 16 14 18

Les compter :

sum(y%%2 == 0)

## [1] 3



# Modes

#### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle
Graphiques de

base base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Les vecteurs contiennent des données de même mode :

• numérique : 1:2, 1L (L pour un entier)

• imaginaire : (1+1i)\*(1-1i) égale 2

logique : TRUE

• caractère : "Bonjour"

• vide : NULL



## Matrices

### Eric Marcon

Installation

Démarrage

## Données

Fonctions

Structures de contrôle

Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

Les matrices ont deux dimensions et contiennent des données de même mode

```
(A <- matrix(1:9, nrow = 3))
```

```
## [,1] [,2] [,3]
## [1,] 1 4 7
## [2,] 2 5 8
## [3,] 3 6 9
```



# Sélectionner dans une matrice

### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

A[1, 2]

## [1] 4

A[, 3]

## [1] 7 8 9



# **Tableaux**

### Eric Marcon

Installation Démarrage

#### Données

Fonctions

Structures de

contrôle Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

### Extension des matrices à plus de deux dimensions

```
A \leftarrow array(1:18, dim = c(3, 3, 2))
A[, , 2]
```

```
## [,1] [,2] [,3]
## [1,] 10 13 16
## [2,] 11 14 17
## [3,] 12 15 18
```



## Sélectionner dans un tableau

### Eric Marcon

Installation

Démarrage

#### Données

Fonctions

Structures de

contrôle Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

### Comme dans une matrice :

A[,,2]

## [,1] [,2] [,3] ## [1,] 10 13 16 ## [2,] 11 14 17 ## [3,] 12 15 18

## Listes

#### Eric Marcon

Installation

### Démarrage

### Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Eléments disparates :

```
(L \leftarrow list(noms = c("X", "Y"), tailles = c(100, 120)))
```

```
## $noms
## [1] "X" "Y"
##
## $tailles
## [1] 100 120
```



## Sélectionner dans une liste

### Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Double crochet ou nom:

L[[2]]

## [1] 100 120

### L\$noms

## [1] "X" "Y"



# Dataframe

### Eric Marcon

Installation

Démarrage

#### Données

Fonctions

Structures de contrôle

Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

### Tableau dont chaque colonne est de mode unique :

```
(df <- data.frame(nom = c("X", "Y"), taille = c(100,
120)))
```

nom	taille
X	100
Υ	120



# Sélectionner dans un dataframe (1)

### Eric Marcon

Installation

Démarrage

#### Données

Fonctions

Structures de contrôle

Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Sélection comme dans une matrice...

df[2,]

	nom	taille
2	Υ	120

... ou comme dans une liste

df\$taille

## [1] 100 120



# Sélectionner dans un dataframe (2)

Eric Marcon

Installation

Démarrage

### Données

Fonctions

Structures de contrôle

Contrôle

Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

### Sélection de lignes en fonction de valeurs

df[df\$taille == 100, ]

nom	taille
Χ	100



#### Eric Marcon

Installation

Démarrage

Données

### Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

## **Fonctions**



# Définition

Eric Marcon

Installation

Démarrage

Données

### Fonctions

Structures de contrôle
Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

R est un langage fonctionnel.

y <- cos(pi)

Une fonction produit une valeur à partir d'arguments.



# Effets de bord

### Eric Marcon

Installation

Démarrage

Données

### Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Une fonction peut avoir des effets de bord :



X

## NULL

plot n'est utilisé que pour ses effets de bord.



# Arguments

#### Eric Marcon

Installation

Démarrage

Données

### Fonctions

Structures de contrôle

Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

Appeler une fonction en nommant tous ses arguments...

```
runif(n = 3, min = 0, max = 1)
```

## [1] 0.10293624 0.02248227 0.91641919

... ou en les passant dans l'ordre :

```
runif(3, 0, 1)
```

## [1] 0.9096517 0.1370688 0.3397247



# Arguments nommés

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

#### Offictions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Bonne pratique : nommer tous les arguments à partir du deuxième:

```
runif(3, min = 0, max = 1)
```

```
## [1] 0.1005619 0.4669571 0.4286474
```



# Valeurs par défaut

Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle
Graphiques de

base de

Tidyverse

Graphiques avec ggplot

Aller plus loin

Voir l'aide de la fonction : ?runif

min et max ont des valeurs par défaut : 0 et 1.

runif(3)

## [1] 0.53216272 0.09179518 0.02105684

# Créer

### Eric Marcon

Installation

Démarrage

Données

### Fonctions

Structures de contrôle
Graphiques de

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

### Syntaxe:

```
puissance <- function(x, r = 1) {
    return(x^r)
}
puissance(1:3, r = 2)</pre>
```

## [1] 1 4 9

Penser vecteur. r est recyclé.

```
puissance(1:3, r = 3:1)
```

## [1] 1 4 3



#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

# Structures de contrôle

# Si / Sinon

### Eric Marcon

Installation

Démarrage

Données

Fonctions

# Structures de contrôle

Graphiques de base

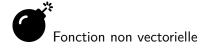
Tidyverse

Graphiques avec ggplot

Aller plus loin

```
est_pair <- function(x) {
   if (x%%2 == 0) {
      return(TRUE)
   } else {
      return(FALSE)
   }
}
est_pair(3)</pre>
```

```
## [1] FALSE
```





# Boucles

```
Eric Marcon
```

Installation

Démarrage

Données

Fonctions

Structures de

contrôle Graphiques de

base Tidyverse

Graphiques avec ggplot

Aller plus loin

```
for (i in 1:3) {
    print(sqrt(i))
}
```

```
## [1] 1
## [1] 1.414214
## [1] 1.732051
```



Seulement si la fonction utilisée n'est pas vectorielle.

```
sqrt(1:3)
```

```
## [1] 1.000000 1.414214 1.732051
```



#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

# Graphiques de base

# plot

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

## Graphiques de base

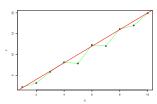
Tidyverse

Graphiques avec ggplot

Aller plus loin

## Graphiques simples:

```
X <- 1:10
Y <- 2 * X + rnorm(length(X))
plot(x = X, y = Y)
lines(x = X, y = Y, col = "green", lty = 2)
abline(a = 0, b = 2, col = "red")</pre>
```





## Classes

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Les objets appartiennent à des classes.

```
Ns <- rlnorm(100)
class(Ns)
```

```
## [1] "numeric"
```

plot est une méthode, déclinée par classe.

```
plot(Ns) # plot.numeric()
```

```
interpretation for
```



## Classes

#### Fric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

library("entropart") Ns <- as.ProbaVector(Ns) class(Ns)

"ProbaVector" [3]

"numeric"

"SpeciesDistribution"

plot(as.ProbaVector(Ns)) # plot.SpeciesDistribution





#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de

contrôle

Graphiques de base

#### Tidyverse

Graphiques avec ggplot

Aller plus loin

# Tidyverse



## Manifeste

#### Eric Marcon

Installation Démarrage

Données

Fonctions

Structures de contrôle
Graphiques de

base

### Tidyverse

Graphiques avec ggplot

Aller plus loin

Univers bien rangé.

Extension de R : ensemble de packages

library("tidyverse")

Manifeste

vignette("manifesto")



# Données

#### Fric Marcon

Installation Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

base

### Tidyverse

Graphiques avec ggplot

Aller plus loin

Autant que possible dans un dataframe.

tibble : dataframe amélioré.

nom	taille	
X	100	
Υ	120	



# Tuyau

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de

Graphiques de

base

### Tidyverse

Graphiques avec ggplot

```
Le résultat d'une fonction est le premier argument de la fonction suivante.
```

```
x <- runif(100, max = 10) %>%
    mean()
x
```

```
## [1] 4.847069
```

```
# ou même
100 %>%
    runif(max = 10) %>%
    mean() %>%
    print() -> x
```

```
## [1] 5.18077
```



# Bagarre (1)

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

#### Tidyverse

Graphiques avec ggplot

Aller plus loin

Data wrangling : lecture des données dans un tibble, sélection des lignes et colonnes, création de colonnes...

```
# Lecture des arbres de la parcelle 6 de Paracou
read csv2("data/Paracou6.csv") %>%
  # Ne garder que les fabaceae
 filter(Family == "Fabaceae") %>%
  # Sélectionner les colonnes espèce et circonférence
  select(spName, CircCorr) %>%
  # Calculer la surface terrière de chaque arbre en m2
  mutate(G = CircCorr^2/4/pi/10000) %>%
  # Grouper par espèce
  group_by(spName) %>%
  # Calculer le nombre de tiges et la surface terrière par ha
  summarize(Abondance = n(), Surface = sum(G)/6.25, .groups = 'drop') %
  # Trier par G/ha décroissant
  arrange(desc(Surface)) ->
 mon tibble
```



# Bagarre (2)

### Eric Marcon

Installation

mstallatio

Démarrage

Données

Fonctions
Structures de

contrôle Graphiques de

base

## Tidyverse

Graphiques avec ggplot

Aller plus loin

## Travail fastidieux:

- prévoir du temps
- capitaliser.

### mon\_tibble

spName	Abondance	Surface
Eperua_falcata	266	5.6776032
Eperua_grandiflora	67	1.5033659
Vouacapoua_americana	91	1.4448320
Dicorynia_guianensis	44	0.7624033
Recordoxylon_speciosum	31	0.4375787
Bocoa_prouacensis	52	0.3083153
Tachigali_melinonii	10	0.3039236
Swartzia_polyphylla	11	0.2366729
Abarema_jupunba	10	0.1656730
Albizia_pedicellaris	3	0.1085323
Andira_coriacea	7	0.0867872
Swartzia_panacoco	6	0.0839838
Parkia_nitida	6	0.0781495
Enterolobium_schomburgkii	4	0.0689622



#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de

contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin

# Graphiques avec ggplot

# ggplot2

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle
Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Package destiné à la création de graphiques.

Respecte la grammaire graphique par couches :

Les données sont obligatoirement un dataframe (un tibble est un dataframe).



# Esthétique

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

Tidyverse

Graphiques avec ggplot

Aller plus loin

L'esthétique désigne ce qui est représenté :

- x et y (ou fill pour un histogramme...)
- transparence, couleur, type de courbe, taille... : voir l'aide de chaque geom\_.

Fonction aes() à plusieurs niveaux :

- argument mapping de ggplot(), hérité par les couches (geom\_)
- ou argument mapping de chaque couche.



## Géométrie

### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

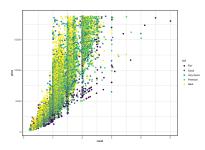
base Tidyverse

Graphiques avec ggplot

Aller plus loin

La géométrie est définie par une fonction geom\_xxx et une esthétique (ce qui est représenté).

```
ggplot(data = diamonds) +
geom_point(mapping = aes(x = carat, y = price, color = cut))
```





# Statistiques (1)

Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

Chaque geom\_ va de pair avec une statistique de transformation des données :

- "identity" pour geom\_point
- "boxplot" pour geom\_boxplot
- 20 statistiques disponibles...



# Statistiques (2)

```
Eric Marcon
```

Installation

Démarrage

Données

Fonctions

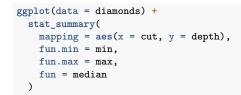
Structures de

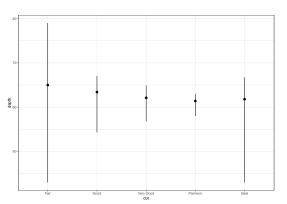
contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot







## **Echelle**

#### Fric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

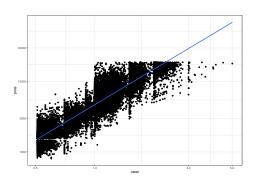
Tidyverse

Graphiques avec ggplot

Aller plus loin

### Transformation de variable.

```
diamonds %>%
  filter(carat > 0.5) %>%
  ggplot(aes(x = carat, y = price)) + geom_point() +
  scale_x_log10() + scale_y_log10() + geom_smooth(method = "lm")
```





# Position (1)

Eric Marcon

Installation

Démarrage

Données

Fonctions

1 Offiction

Structures de contrôle
Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

La position définit l'emplacement des objets sur le graphique.

- "identity" en général,
- "stack" empile les catégories dans un histogramme,
- "jitter" déplace aléatoirement les points dans un geom\_point pour éviter les superpositions.



# Position (2)

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

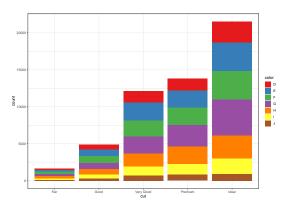
Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot







# Coordonnées (1)

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de

contrôle Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

## Système de coordonnées :

- coord\_flip() intervertit x et y,
- coord\_polar() : coordonnées polaires,
- coord\_trans() transforme l'affichage des coordonnées (mais pas les données comme scale\_),
- etc.

Exemple: tracer la carte des wapas de la parcelle 6.



# Coordonnées (2)

```
Eric Marcon
```

Installation

Démarrage

Données

Fonctions

Structures de

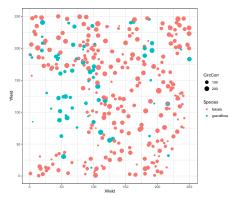
contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot







## **Facettes**

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

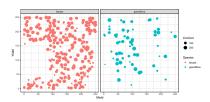
Tidyverse

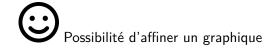
Graphiques avec ggplot

Aller plus loin

## Présente plusieurs aspects du même graphique:

P6Map + facet\_wrap(~Species)







#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot

Aller plus loin



# Rédiger avec RMarkdown

Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de

contrôle Graphiques de

Tidyverse

Graphiques avec ggplot

Aller plus loin

Plutôt qu'un code commenté, un texte avec du code.

Tricot: production de documents HTML ou PDF.

Rédaction d'articles, de mémoires, de diaporama.

Reproductibilité : le projet contient les données, le code, le texte et le modèle de mise en forme.

 ${\sf Galerie:https://ericmarcon.github.io/memoiR/}$ 



## Contrôle de source

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de

contrôle Graphiques de

base

Tidyverse

Graphiques avec ggplot

Aller plus loin

git (contrôle de source) et GitHub (plateforme web) pour :

- tracer les versions d'un projet,
- collaborer,
- tester le code automatiquement,
- tricoter automatiquement.

 ${\sf Exemple: https://github.com/EricMarcon/travailleR}$ 



## Sites web etc.

#### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de

Tidyverse

Graphiques avec ggplot

Aller plus loin

Il existe des packages pour tout.

## Exemples:

- Site web: https://ericmarcon.github.io/fr/,
- Application Shiny : https://vac-lshtm.shinyapps.io/ncov\_tracker/,
- TP en ligne: https://eric-marcon.shinyapps.io/TP-Biodiversite/.



### Eric Marcon

Installation

Démarrage

Données

Fonctions

Structures de contrôle

Graphiques de base

Tidyverse

Graphiques avec ggplot