Renaudpack2 - tuto

Contents

1	Inti	Introduction			
	1.1	Pour l'installer	3		
	1.2	Pour le mettre à jour	3		
	1.3	Pour le charger	3		
2	Les	fonctions graphiques pour les analyses univari?es	4		
	2.1	Un facteur et des barres d'erreur avec la fonction barres.plot	4		
	2.2	Pour rajouter les lettres des post-hoc avec la fonction anovLetters	6		
	2.3	Deux facteurs et des barres d'erreur avec la fonction barres.plot.beside	7		
	2.4	Faire un tableau r?capitulatif avec Tableau_recap	11		
	2.5	Des s?ries temporelles avec Time.factor.plot	13		
	2.6	Des jolies couleurs avec les fonctions Couleur_continue et modif_coul	13		
	2.7	Des diagrammes en barres mais avec des points avec point.plot	17		
3	Fon	actions graphiques pour les analyses multivar?es	19		
	3.1	La fonction multivar.polyg	19		
	3.2	La fonction MultiDyn	25		
	3.3	La fonction better_arrows	26		
	3.4	La fonction label.corV2	27		
	3.5	La fonction label.corDCA	31		
	3.6	La fonction label.corNMDS	32		
4	Pou	Pour de la manipulation de données			
	4.1	La fonction BBtransf	33		
	4.2	La fonction Classes_def	33		
	4.3	La fonction combin.tab	33		
	4 4	La fonction combin tabVO	33		

5	Pou	Pour des t?ches tr?s sp?cifiques:		
	5.1	Autour de l'indice de l'integrit? de la structure des communaut?s	33	
	5.2	La fonction ComStructIndices	33	
	5.3	La fonction structure.plot	33	
	5.4	La fonction structure.plotV2	33	
	5.5	La fonction "	33	
$D\epsilon$	ernièr	$re\ mise\ a\ jour:\ 25/02/2022$		

1 Introduction

Il s'agit d'un tuto pour mes fonctions perso, réunnies dans le package au doux sobriquet de Renaudpack2. Pour quoi Renaupack2 égoet pas un nom un peu moins égocentré? Pour rappeller que c'est un package personnel, qui regroupe mes fonctions perso.. Alors oui, je le partage volontier, mais non, je ne veux pas que ça devienne un package classique du CRAN dont je devrais assurer la maintenance et le service après-vente...

Il s'agit pour la plupart de fonctions écrites en R base et point de ggplot2 ou de dplyr dedans.. Je suis de la vieille école et suis relativement à l'aise avec les bricolages de R base. Les grammaires spécifiques à ces nouveaux écosystèmes de package ggplot2 ou de dplyr, je ne me les suis jamais vraiment approprié et j'ai l'impression d'aller plus vite avec base.. C'est donc fort probable que des solutions plus fancy existent dans ces nouveaux packages, pour faire la même chose que ce que je propose ici..

Depuis peu l'installation est possible via Github, c'est ultra simple en utilisant une fonction du package devtools, qu'il faut donc potentiellement installer si ce n'est déjà fait. Mais depuis peu également, un autre package rtools dont dépends devtools, doit être installé comme un plug-in à part, si vous ne l'avez jamais fais, téléchargez et installer rtools à partir du lien suivant : https://cran.r-project.org/bin/windows/Rtools/

1.1 Pour l'installer

```
install.packages("devtools")
#? ne faire que si vous n'avez jamais installé devtools
library(devtools)
devtools::install_github("RenaudJau/Renaudpack2")
```

1.2 Pour le mettre à jour

```
install.packages("devtools")
#? ne faire que si vous n'avez jamais install? devtools
library(devtools)
devtools::update_packages("Renaudpack2")
```

1.3 Pour le charger

Comme n'importe quel package, avec library:

library(Renaudpack2)

2 Les fonctions graphiques pour les analyses univari?es

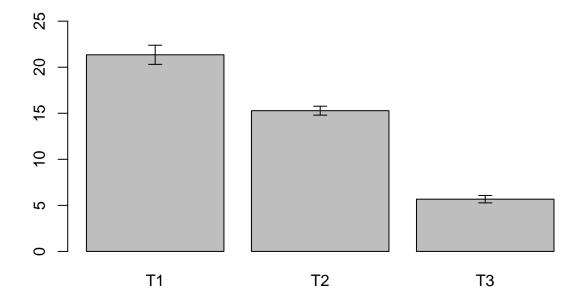
Pour les besoins de ce tuto, on créé des fausses données :

```
biomasse<-c(rnorm(35,20,5),rnorm(35,15,3),rnorm(35,5,3))
traitement<-factor(rep(c("T1","T2","T3"),each=35))
```

2.1 Un facteur et des barres d'erreur avec la fonction barres.plot

Le graphique de base :

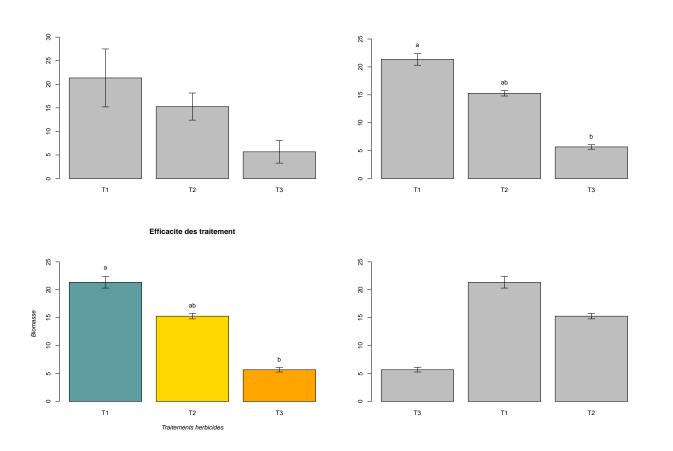
barres.plot(biomasse,traitement)



Il y a plusieurs choses que l'on peut modifier pour personaliser son graphique :

- changer le type d'erreur représentée
- ajouter des lettres d'un test post-hoc
- changer des param?tres classiques de la fonction barplot
- changer l'ordre des modalit?s (?a ?a n'est pas sp?cifique ? la fonction barres.plot)
- on peut aussi modifier l'axe des abscisses (cf les arguments las.x, cex.x et labels.x de la fonction barres.plot.beside expliquée plus tard.

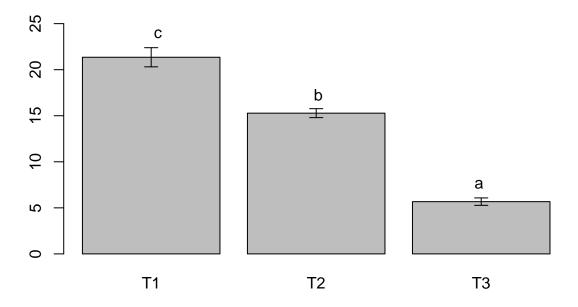
```
par(mfrow=c(2,2)) #pour avoir 4 graphiques sur une m?me fen?tre
barres.plot(biomasse,traitement,ecart=sd)
#erreur standard par d?faut, ici ?cart-type
barres.plot(biomasse,traitement,lettres=c("a","ab","b"))
barres.plot(biomasse,traitement,lettres=c("a","ab","b")),
col=c("cadetblue","gold","orange"),
xlab="Traitements herbicides",ylab="Biomasse",font.lab=3,
main="Efficacite des traitement")
traitement2<-factor(traitement,levels=c("T3","T1","T2"))
barres.plot(biomasse,traitement2)</pre>
```



2.2 Pour rajouter les lettres des post-hoc avec la fonction anovLetters

Si vous ne voulez pas vous embêter à calculer vous même les lettres des tests post-hoc, et si et uniquement si vous ?tes dans les conditions d'utilisation d'une ANOVA ? un facteur, alors on peut faire ?a :

```
anov<-aov(biomasse~traitement)</pre>
summary(anov)
##
                 Df Sum Sq Mean Sq F value Pr(>F)
## traitement
                  2
                      4373 2186.6
                                       126.8 <2e-16 ***
## Residuals
                102
                       1758
                               17.2
## Signif. codes: 0 '*** 0.001 '** 0.01 '* 0.05 '.' 0.1 ' ' 1
OK, il y a un effet significatif, du coup anovLetters renvoie la liste des lettres...
lettresPH<-anovLetters(VAR = biomasse,FAC = traitement)</pre>
lettresPH
## [1] "
          c" " b " " a "
...que l'on peut utiliser directement sur barres.plot :
```



barres.plot(biomasse,traitement,lettres = lettresPH)

^{*}Note : possibilité de changer le seuil alpha en utilisant l'argument ALPHA.

2.3 Deux facteurs et des barres d'erreur avec la fonction barres.plot.beside

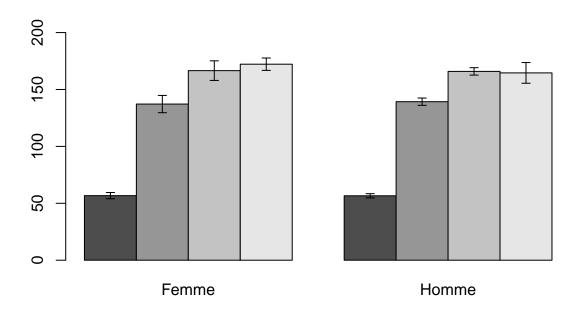
Pour les besoins de ce tuto, on cr?? des fausses donn?es, avec 2 facteurs, l'?ge et le sexe, et une variable, la taille :

En gros ?a donne ?a :

Age	Sexe	Taille
2	Homme	59.85671
2	Femme	61.19198
2	Homme	57.67089
2	Femme	63.84244
2	Homme	57.70931
2	Femme	48.74355

Le graphique de base :

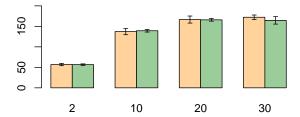
```
barres.plot.beside(Taille,Sexe,Age)
```



On peut aussi modifier quelques petits trucs :

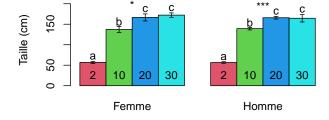
• si on inverse les facteurs, avec des couleurs

```
barres.plot.beside(Taille,Age,Sexe,col=c("burlywood1","darkseagreen3"))
```



• avec les annotations qui vont bien

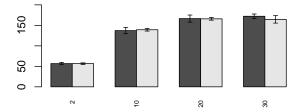
Attention, les lettres et ?toiles sont ? d?finir soit m?me, c-?-d apr?s tests statistiques, ici pour l'exemple, c'est de l'al?atoire...), et avec des couleurs simples (mais moches...)



On peut aussi modifier l'axe des abscisses :

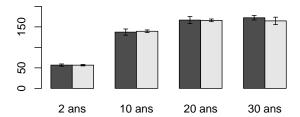
• l'orientation avec las.x et la taille avec cex.x :

barres.plot.beside(Taille,Age,Sexe,las.x = 2, cex.x = 0.6)



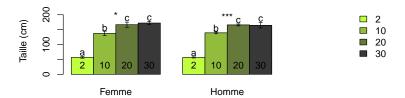
• les étiquettes avec labels.x:

barres.plot.beside(Taille,Age,Sexe,labels.x = c("2 ans","10 ans","20 ans","30 ans"))

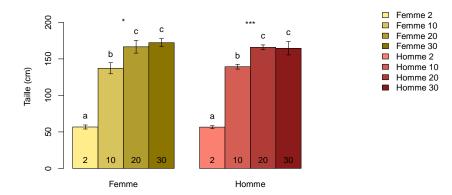


Attention, changer les noms ne veut pas dire qu'on change l'ordre des facteurs.. ne pas écrire n'importe quoi..

• avec des couleurs choisies avec la fonction colorRampPalette et une l?gende



• avec des couleurs pour chaque combinaison de modalit?s



2.4 Faire un tableau r?capitulatif avec Tableau_recap

Cette fonction n'est pas graphique, elle donne les informations de barres.plot sous forme d'un tableau. Elle a peu d'utilit? en dehors d'un document Rmarkdown.

```
Tableau_recap(VAR = biomasse, FAC = traitement, ROUND = 2)
```

```
##
      Modalites Nombre Moyennes Erreur
## T1
             T1
                     35
                           21.35
## T2
             T2
                     35
                           15.28
                                   0.49
## T3
             Т3
                     35
                           5.67
                                   0.40
```

Pour que ce soit plus joli, il est cosneill? d'utiliser la fonction kabledu package knitr (pour l'installer : install.packages("knitr")).

```
kable(Tableau_recap(VAR = biomasse, FAC = traitement, ROUND = 2))
```

	Modalites	Nombre	Moyennes	Erreur
T1	T1	35	21.35	1.04
T2	T2	35	15.28	0.49
Т3	Т3	35	5.67	0.40

Si on fait un test posthoc, on peut rajouter les lettres dans le tableau :

```
anov<-aov(biomasse~traitement)
summary(anov)</pre>
```

```
## Df Sum Sq Mean Sq F value Pr(>F)
## traitement 2 4373 2186.6 126.8 <2e-16 ***
## Residuals 102 1758 17.2
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1</pre>
```

```
TukeyHSD(anov)
     Tukey multiple comparisons of means
##
##
       95% family-wise confidence level
##
## Fit: aov(formula = biomasse ~ traitement)
##
## $traitement
##
               diff
                           lwr
                                      upr p adj
## T2-T1 -6.069363 -8.429958 -3.708769 1e-07
## T3-T1 -15.675764 -18.036359 -13.315170 0e+00
## T3-T2 -9.606401 -11.966995 -7.245806 0e+00
#les lettres sont donc a, b et c :
kable(Tableau_recap(VAR = biomasse, FAC = traitement, ROUND = 2,
              LETTRES = c("a","b","c")))
```

	Modalites	Nombre	Moyennes	Erreur	Lettres
T1	T1	35	21.35	1.04	a
T2	T2	35	15.28	0.49	b
T3	Т3	35	5.67	0.40	\mathbf{c}

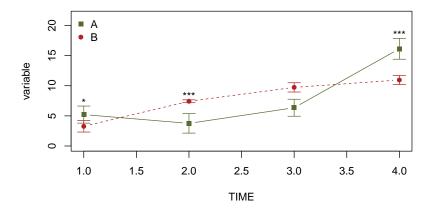
2.5 Des s?ries temporelles avec Time.factor.plot

Pour les besoins de ce tuto, on cr?? des fausses donn?es :

En gros ?a donne ?a :

$\overline{ ext{tim}}$	fac	vari
1	A	8.044418
1	A	3.831582
1	A	3.751536
1	В	5.088431
1	В	1.922498
1	В	2.769196

Utilisation de la fonction : Pourquoi ce message d'erreur? Bonne question.. ? r?soudre..



2.6 Des jolies couleurs avec les fonctions Couleur_continue et modif_coul

Couleur_continue permet de transformer une variable en palette de couleur, utile pour ajouter une 'dimension'? un plot en 2D

Par exemple, si on a des donn?es (fausses ici..) avec la couverture v?g?tale en fonction de l'altitude, et une information sur la taille des plantes :

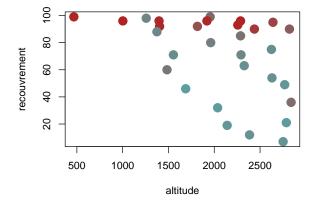
taille	recouvrement	altitude
89.2	99	467
18.0	98	1257
40.1	99	1955
90.6	96	2287
86.0	96	1920
83.9	96	1397

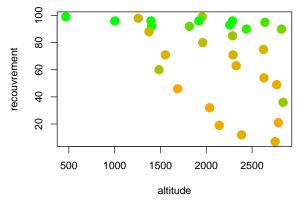
On peut cr?er une couleur pour chaque valeur de taille (ici avec des couleurs plus froides pour les tailles les plus petites) :

```
coul_taille<-Couleur_continue(VAR = taille,COLORS = c("cadetblue","firebrick"))
coul_taille</pre>
```

```
## [1] "#B02424" "#6F8586" "#836768" "#B22222" "#AD2828" "#AC2A2A" "#B22222" ## [8] "#AD2929" "#925151" "#B02424" "#A33838" "#9A4445" "#8B5B5B" "#7C7172" ## [15] "#737F80" "#6D8789" "#688F90" "#689092" "#689092" "#649597" "#806C6D" ## [22] "#619A9C" "#659495" "#669294" "#639698" "#619A9C" "#5F9EA0" "#787778" ## [29] "#698D8F" "#6D888A"
```

Ces couleurs (en code h?xad?cimal) peuvent ?tre utilis?es dans un graphique, avec la possibilit? de changer les couleurs comme on veut :

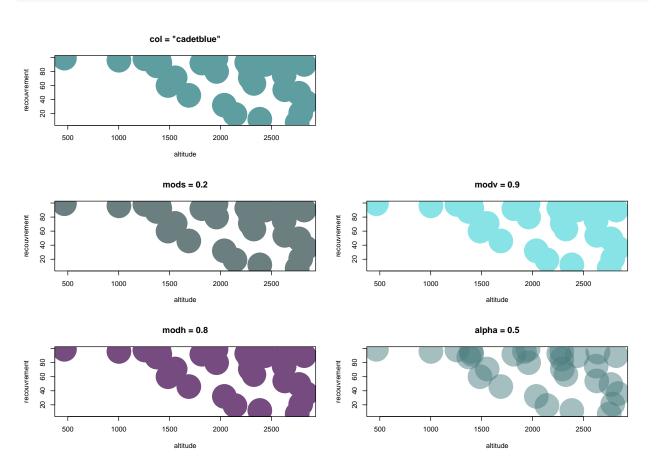




La fonction modif_coul permet de modifier une couleur existante :

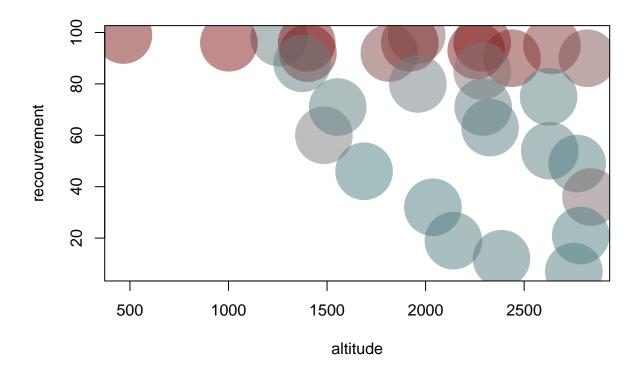
Par exemple, en partant de la couleur cadetblue (qui est un bleu pastel), on peut :

- modifier la saturation, sous 0.5 ?a d?sature, au dessus, ?a sature
- modifier la brillance, sous 0.5 ?a fonce, au dessus, ?a ?claircit
- modifier la teinte, ?a tourne en rond..
- modifier la transparence sous 1, ?a devient transparent



On peut aussi combiner les deux fonctions :

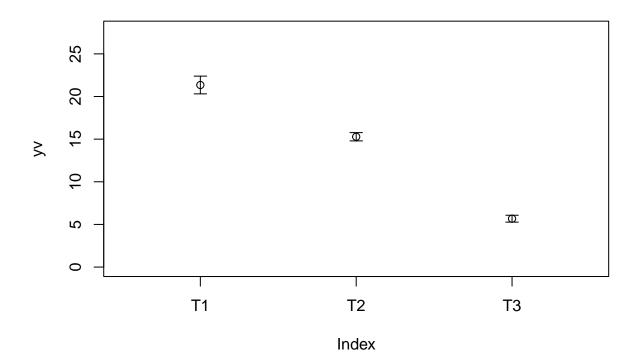
```
coul_taille_transp <- sapply(coul_taille, function(x) modif_coul(x, alpha = 0.5))
plot(x = altitude, y = recouvrement, pch=16, cex=8, col=coul_taille_transp)</pre>
```



2.7 Des diagrammes en barres mais avec des points avec point.plot

Quasi identique ? l'utilisation de barres.plot.. Le graphique de base :

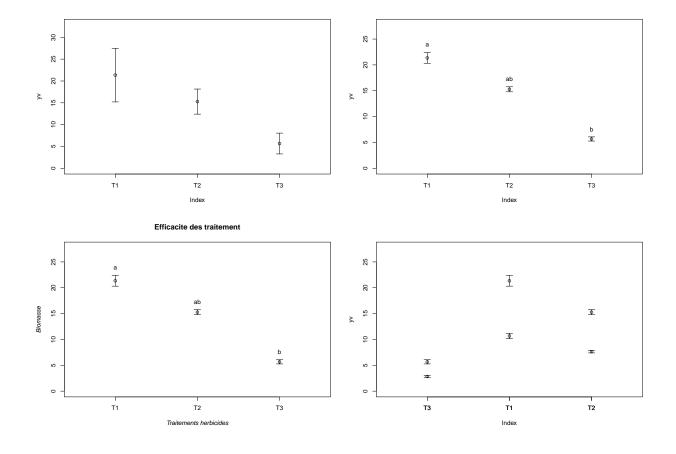
point.plot(biomasse,traitement)



Il y a plusieurs choses que l'on peut modifier pour personaliser son graphique :

- changer le type d'erreur repr?sent?e
- ajouter des lettres d'un test post-hoc
- changer des param?tres classiques de la fonction barplot
- changer l'ordre des modalit?s (?a ?a n'est pas sp?cifique ? la fonction point.plot) et surajouter sur un graphique

```
par(mfrow=c(2,2)) #pour avoir 4 graphiques sur une m?me fen?tre
point.plot(biomasse,traitement,ecart=sd)
#erreur standard par d?faut, ici ?cart-type
point.plot(biomasse,traitement,lettres=c("a","ab","b"))
point.plot(biomasse,traitement,lettres=c("a","ab","b"),
xlab="Traitements herbicides",ylab="Biomasse",font.lab=3,
main="Efficacite des traitement")
traitement2<-factor(traitement,levels=c("T3","T1","T2"))
point.plot(biomasse,traitement2)
point.plot(biomasse/2,traitement2, add = TRUE)</pre>
```



3 Fonctions graphiques pour les analyses multivar?es

3.1 La fonction multivar.polyg

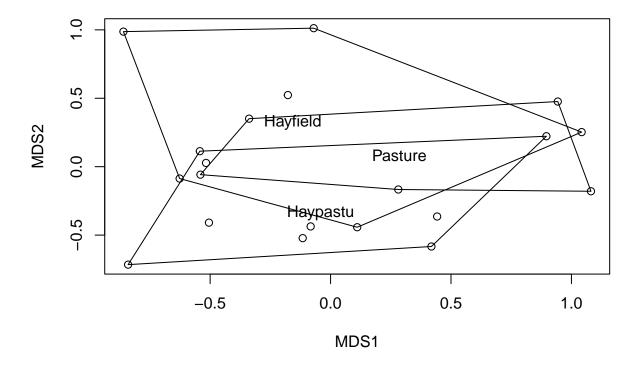
La fonction multivar.polyg permet de regrouper des points dans une analyse en faisant un polygone autour. On prends

Il faut:

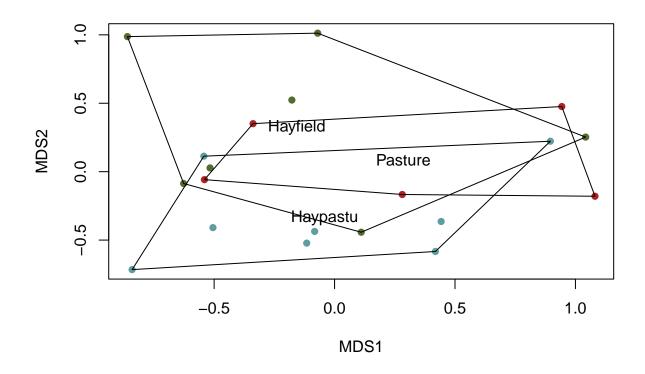
- ANAcoo coordonnées des points d'une analyse multivariées (__\$points pour une NMDS, __\$li pour une AFC, etc.)
- FAC facteur qui permettra de séparer les polygones

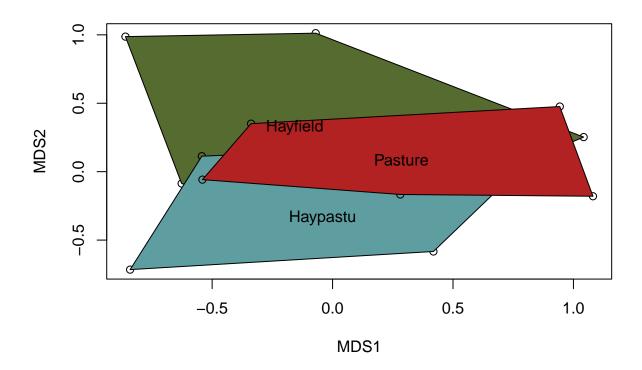
```
library(vegan)
## Le chargement a nécessité le package : permute
## Le chargement a nécessité le package : lattice
## This is vegan 2.6-4
data("dune")
data("dune.env")
NMDS <- metaMDS(dune)</pre>
## Run 0 stress 0.1192678
## Run 1 stress 0.1192678
## ... Procrustes: rmse 2.330208e-05 max resid 7.002553e-05
## ... Similar to previous best
## Run 2 stress 0.1808911
## Run 3 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 0.02027093 max resid 0.0649643
## Run 4 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 7.0794e-06 max resid 1.995902e-05
## ... Similar to previous best
## Run 5 stress 0.1183186
## ... Procrustes: rmse 1.901128e-06 max resid 5.040964e-06
## ... Similar to previous best
## Run 6 stress 0.1192679
## Run 7 stress 0.1192678
## Run 8 stress 0.192224
## Run 9 stress 0.1192679
## Run 10 stress 0.2075713
## Run 11 stress 0.1183186
## ... Procrustes: rmse 3.124662e-05 max resid 8.211294e-05
## ... Similar to previous best
## Run 12 stress 0.1183186
## ... New best solution
## ... Procrustes: rmse 3.838469e-06 max resid 1.349485e-05
## ... Similar to previous best
```

```
## Run 13 stress 0.1192678
## Run 14 stress 0.1808911
## Run 15 stress 0.1183186
## ... Procrustes: rmse 9.725646e-06 max resid 2.908327e-05
## Run 16 stress 0.1192679
## Run 17 stress 0.1192678
## Run 18 stress 0.1808912
## Run 19 stress 0.1192678
## Run 20 stress 0.1808911
## *** Best solution repeated 2 times
multivar.polyg(ANAcoo = NMDS$points, FAC = dune.env$Use)
```

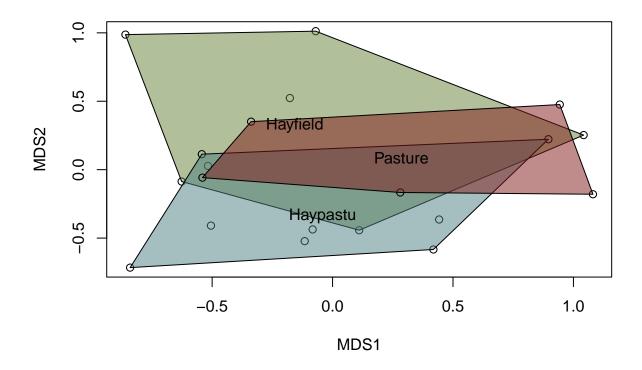


Avec d'autres arguments si besoin :

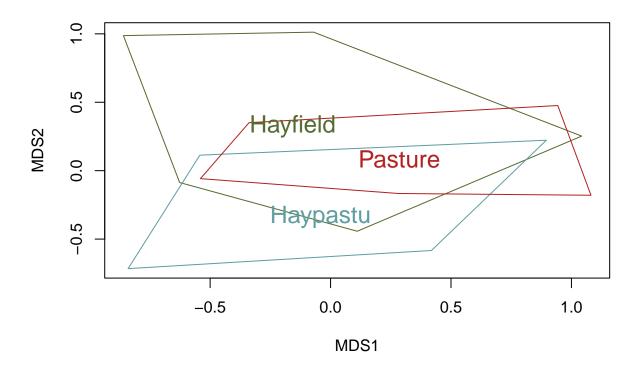




 $\label{lem:modif_coul} \mbox{Mieux vaux utiliser la fonction } \mbox{{\tt modif_coul}\ pour\ ajouter\ de\ la\ transparence}:$



Et encore :



Possibilité aussi avec les arguments lab de ne pas afficher les étiquettes des modalités, new de ne pas faire ça sur un nouveau graphe mais sur un graphe existant, et sep de faire une nouvelle fenêtre pour chaque modalité du facteur.

3.2 La fonction MultiDyn

3.3 La fonction better_arrows

3.4 La fonction label.corV2

La fonction label.corV2 permet de calculer les corrélations des espèces avec les axes et d'afficher seulement les plus corrélées.

Il y a plusieurs arguments obligatoires:

- COO_ESP Coordonnées des espèces suites à une ordination, pour les deux premiers axes, après dudi.pca et dudi.coa c'est ANALYSE\$co[,1:2], après decorana c'est ANALYSE\$cproj[,1:2], après metaMDS c'est ANALYSE\$species[,1:2], etc.
- COO_REL Coordonnées des relevés suites à une ordination, pour les deux premiers axes, après dudi.pca et dudi.coa c'est ANALYSE\$li[,1:2], après decorana c'est ANALYSE\$rproj[,1:2], après metaMDS c'est ANALYSE\$points[,1:2], etc.
- RELEVES les relevés avant servi à faire l'analyse multivariée
- METHOD Méthode de choix des espèces à afficher, soit "P" (défaut) les espèces dont le p du test de spearman sont inférieures à la valeur de P sur au moins un axe sont affichées, soit "RHO" les espèces dont le Rho du test de spearman sont supérieur à RHO sur au moins un axe sont affichées, soit "N_base_P" les N espèces dont le P sont les plus faibles sont affichées, soit "N_base_RHO" les N espèces dont les RHO sont les plus fort sont affichées.

Selon la méthode choisie, vous devrez aussi renseigner :

- P la p value maximum pour être affichée
- RHO le Rho minimum pour être affichée
- N le nombre d'espèces à afficher, attention, si des ex-aequo, pour ne pas surcharger, il y a moins d'espèces affichées que le N choisi

Et éventuellement :

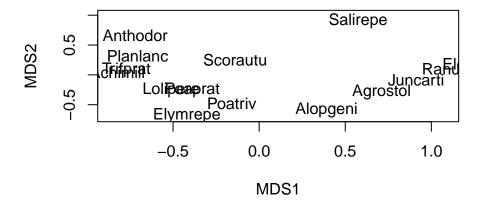
- COEF un coef multiplicateur (pour exploser ou non la dispersion dans le plan)
- ADD par defaut, créer un graphe seul, si add=T, les noms d'especes sont ajouté au graphe déjà existant
- ... d'autres arguments de la fonction text

Pour l'exemple, on récupère les données dune du package vegan et on fait une NMDS :

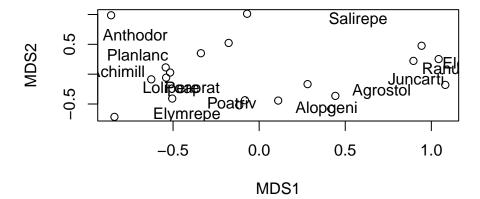
```
library(vegan)
data("dune")
NMDS <- metaMDS(dune)</pre>
```

Usage de la fonction :

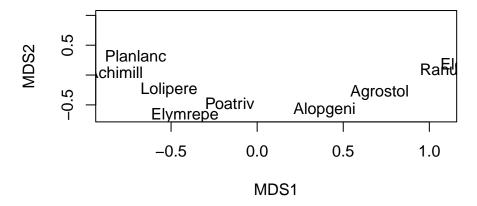
• En fonction du p, et sans graphe existant :



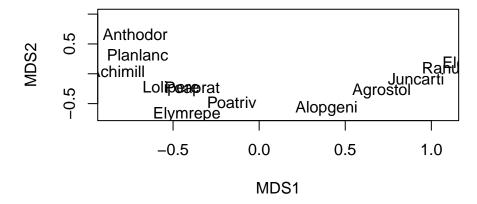
• En fonction du rho et avec graphe existant :



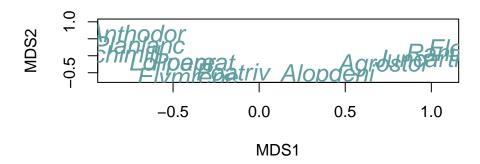
- Sans graphe précédent, par un nombre d'espèces en fonction du ${\bf p}$:



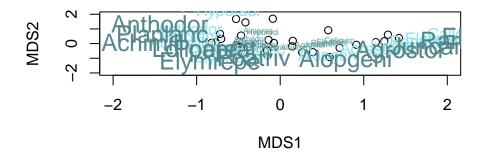
• Sans graphe précédent, par un nombre d'espèces en fonction du rho :



• Sans graphe précédent, par un nombre d'espèces en fonction du rho et en utilisant tout un tas d'autres arguments de la fonction text:



• Et pour voir les effets de l'argument COEF :



3.5 La fonction label.corDCA

3.6 La fonction label.corNMDS

4 Pour de la manipulation de données

- 4.1 La fonction BBtransf
- 4.2 La fonction Classes_def
- 4.3 La fonction combin.tab
- 4.4 La fonction combin.tabV0
- 5 Pour des t?ches tr?s sp?cifiques :
- 5.1 Autour de l'indice de l'integrit? de la structure des communaut?s
- 5.2 La fonction ComStructIndices
- 5.3 La fonction structure.plot
- 5.4 La fonction structure.plotV2
- 5.5 La fonction "

"

"

"