

## Cvičení 4

1. Importujte dataset o popisující rybí společenstva v řece Doubs. Budete potřebovat pouze data o druhovém složení.
2. Proveďte souhrn dat a zkontrolujte, že se načetla správně
3. Spočítejte na datech PCA. Data obsahují hodně nul. Použijte proto Hellingerovu standardizaci.
4. Zobrazte summary analýzy a množství variability vysvětlené osami
5. Zhodnoťte významnost os. Použijte k tomu screeplot.
6. Prostudujte si funkci ordiplot a její použití – jednak v popisu dole a jednak v R dokumentaci.
7. Nakreslete ordinační diagram pro první a druhou osu a pro první a třetí osu; zobrazte na ordinačním diagramu druhy a lokality. Nastavte škálování ordinačních skóre na korelace mezi druhy.  
*Odtud pracujte samostatně ve skupinách*
8. Importujte dataset popisující rostlinná společenstva v Bílých Karpatech. Budete potřebovat data o proměnných prostředí
9. Proveďte souhrn dat a zkontrolujte, že se načetla správně
10. Spočítejte na datech PCA. Vzhledem k tomu, že jde o měřené proměnné, proveďte standardizaci po proměnných. Druhá proměnná je kategoriální – tu musíte z analýzy vyřadit
11. Zhodnoťte významnost os. Použijte k tomu screeplot.
12. Nakreslete ordinační diagram pro významné osy; zobrazte na ordinačním diagramu proměnné pomocí šipek a názvů. Lokality zobrazte pomocí bodů. Škálování nastavte symetrické.
13. Zkuste zjistit, jak se výsledky analýzy změní, když provedete standardizaci po vzorcích.

vegan:

**rda()** Analýza hlavních komponent (PCA), Redundanční analýza (RDA)

PCA je možné zadat pomocí `rda(matice druhů/proměnných)`  
`scale = T` provede standardizaci po druzích (proměnných)

**plot.rda()**, nakreslí ordinační diagram, **scores()** extrahuje skóre na ordinačních osách (používají velmi podobný set parametrů)

`choices` výběr ordinačních os

`display` specifikuje typ skóre

= "sites" skóre vzorků

= "species" skóre druhů (proměnných)

`scaling` specifikuje, která skóre se mají škálovat podle eigenvalues

= 1 nebo "sites" vzorky; jejich vzdálenosti pak odpovídají nepodobnostem

= 2 nebo "species" druhy (proměnné); úhly mezi šipkami pak odpovídají korelacím/kovariancím mezi nimi

= 3 nebo "symmetric" symetrické škálování vzorků i druhů odmocninou z eigenvalues

`xlab`, `ylab` doporučuju používat k znázornění množství vysvětlené variability, např. `xlab = "PC1 (30.5%)"`

**ordiplot()** alternativní (a lepší) možnost kreslení ordinačních diagramů. Většina parametrů je stejná jako ve `scores()`. Nejlepší zkúsob kreslení je použít `ordiplot` s parametrem `type='n'` a výsledek uložit do objektu. tj.

**plot.obj<-ordiplot(pca.obj.,type='n', další parametry)**. Výsledkem je prázdný ordinační diagram do kterého je možné přidávat objekty pomocí funkcí text nebo points.

**text(ordiplot), points(ordiplot)**

what co se bude kreslit (odpovídá par. display ve scores)

arrows = T skóre jsou znázorněna pomocí šipek (vhodné pro druhy nebo proměnné v lineárních ordinacích)

length důležitý parametr, který ovlivňuje velikost hrotu šipek (viz základní funkce arrows())

**eigenvals()** extrahuje eigenvalues ordinačních os

**screeplot()** nakreslí screeplot eigenvalues

bstick = T nakreslí do barplotu očekávání dle broken-stick modelu

abline(h=mean(eigenvals(ordination.object)), lty=2) přidá do orázku čárkovanou čáru znázorňující průměrné eigenvalue přes všechny osy (Kaiser-Guttman kritérium)