

STAT660/FES758b
Multivariate Statistics
Homework #6 OPTION A: Ordination
Due : Monday, 4/22/19 Submit on CANVAS by midnight

For this assignment, you can either use your own data or the data described below. Use any combination of R/SAS/MINITAB/SPSS/STATA that you like. Whichever data you choose, do the following:

- 1) Fit Correspondence Analysis to your data.
- 2) Discuss the inertia, make a two dimensional plot of the first two CA directions.
- 3) Comment on whether or not there is any evidence of 'data snaking' in higher dimensional space.
- 4) In a few sentences, describe what you conclude from your plot.
- 5) Perform Multidimensional Scaling (metric or non-metric) for 1, 2, and 3 dimensions.
- 6) Discuss the stress (or SStress) of each dimensional solution. Make a scree plot if you're able.
- 7) Make a two dimensional plot of your results.
- 8) If possible, overlay some other variables to interpret your ordination axes.
- 9) **BONUS** – try canonical correspondence analysis, or calculate p-values for the overlaid additional variables.

.....

Loaner Data: choose ONE

Cereal.attitudes.csv : Marketing Survey Attitudes toward Cereals

- 8 Cereals
- 11 Questions (come back to, tastes nice, popular with all the family, very easy to digest, nourishing, natural flavor, reasonably priced, a lot of food value, stays crispy in milk, helps to keep you fit, fun for children to eat)
- Values are percent of respondents who had a favorable response for a particular cereal for that particular question.

T. K. Chakrapani and A. S. C. Ehrenberg, "An Alternative to Factor Analysis in Marketing Research Part 2: Between Group Analysis", PMRS Journal, Vol. 1, Issue 2, October 1981, pp. 32-38.

R code to get you started :

```
#get the data
cereal <-
read.csv("http://reuningscherer.net/stat660/data/cereal.attitudes.csv")
```

Wisconsin.Forest.csv : Relative abundance of 14 species was measured on 10 plots. Plots were ordered from pioneer (early stage) to climax (late stage). The final column contains that stage of the forest on a scale from 1 to 10.

Peet & Loucks (1977)

R code to get you started :

```
#get the data
forest <-
read.csv("http://reuningscherer.net/stat660/data/Wisconsin.Forest.csv")
rownames(forest)=forest[,1]
forestenv=matrix(forest[,17],ncol=1)
rownames(forestenv)=forest[,1]
colnames(forestenv)=c("Stage")
forest=forest[,-c(1,17)]
forestenv=data.frame(forestenv)
```