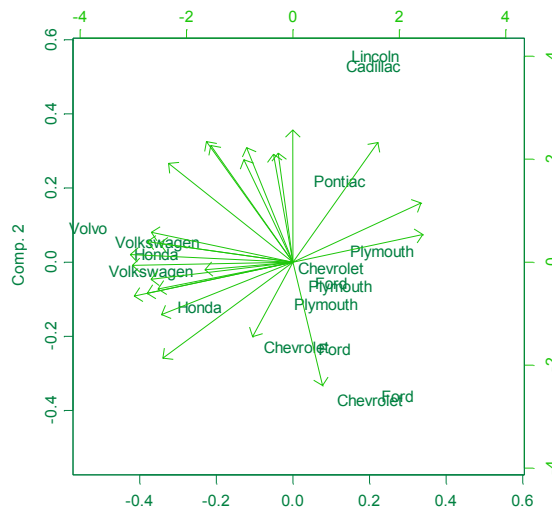
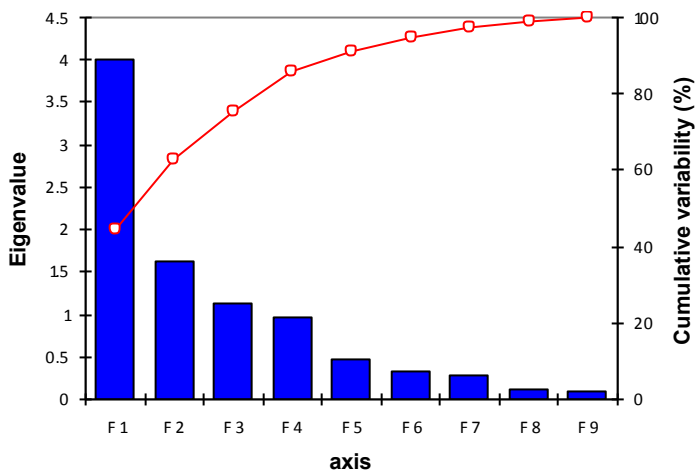


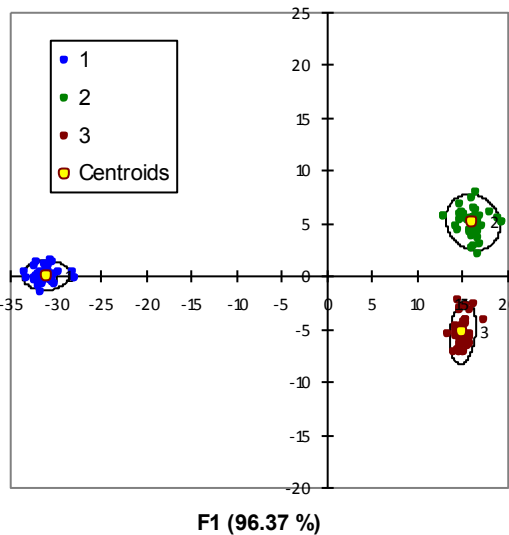
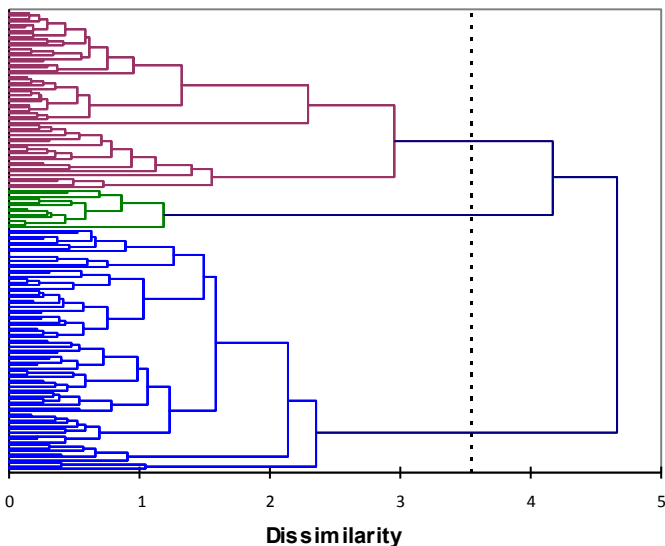
# Data Analysis School

Scree plot



Observations (axes F1 and F2: 100.00 %)

Dendrogram



## Multivariate Data Analysis School

Creascience invites you to the next edition of the Multivariate Data Analysis School that will be held in the heart of downtown Montréal, Canada.

### Practical Details

The course starts at 8h30 and ends at 16h30.

Course materials and coffee-breaks are included in the registration fee.

### Course Overview

Multivariate statistics provide the ability to analyze complex sets of data. They provide a solution for analyzing datasets where there are many independent and possible dependent variables which are correlated to each other to varying degrees. Generally speaking multivariate methods may be used to:

- To plot large sets of data
- To identify groups of inter-related variables
- To reduce the number of variables (dimensionality)
- To predict group membership from a set of variables
- To detect natural groupings in data sets
- To detect multivariate outliers
- And much more...

The ready availability of software application programs which can handle the complexity of large multivariate data sets has increased and popularized the use of multivariate statistics.

### Software and Datasets

- Participants may use their own data for exercises
- Participants are invited to use their own software.

Featured software include R, S-Plus, SAS, SPSS, Statistica, Minitab, S-Plus, JMP and XLStat\*. The workshop also provides an opportunity to compare these packages.

\*If you are using another software package and want to attend the training session, please contact us.

### Testimonials from Past Participants

Take a look at testimonials from past participants: <http://training.creascience.com/reviews.php>

### Contact Information

For more information, please contact Natalie Rodrigue at 514 840 9220 ext 112 (Natalie.[Rodrigue@creascience.com](mailto:Natalie.Rodrigue@creascience.com)).

## Course Curriculum

The multivariate data analysis school will cover the following topics.

### Introduction to Multivariate Methods

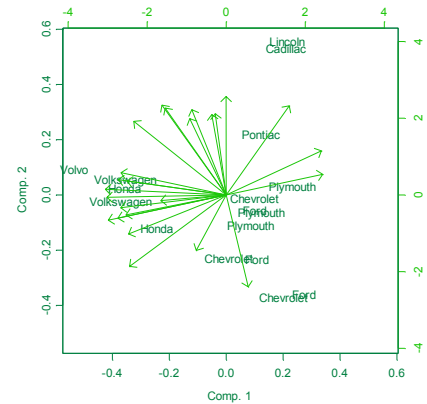
- Why use multivariate methods?
- What information do they convey compared to classical methods?
- How do they work?
- What methods are available? Use of results, interpretation and reporting

### Basic Concepts in Multivariate Analysis

- The notions of variables and objects
- Calculation of the distance between objects : Euclidian, Manhattan
- Types of variables
- Distance between variables : variance, covariance
- Concept of independence
- Hands-on exercises and group discussions

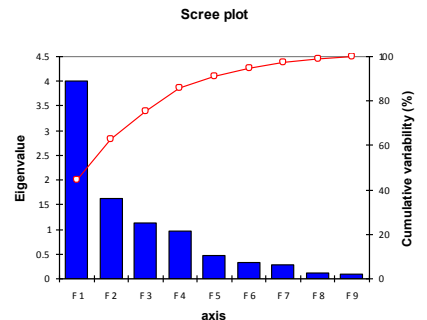
### Principal Component Analysis (PCA)

- Problem, context of use and objectives
- Historical background
- Principle
- Steps involved in principal component analysis
- Determining the Number of components to retain
- Graphical tools: Scree plots
- Statistical indicators
- Interpretation of principal components
- Loadings of variables
- Coordinate of objects
- Correlation circle
- Biplots
- Use of principal components
- Applications
- Hands-on exercises and group discussions



### Factor Analysis

- Problem, context of use and objectives
- Principle
- Latent variables
- Extraction of factors
- Rotations
- Interpretation of factors
- Graphical tools
- Use of factors
- Applications
- Hands-on exercises and group discussions

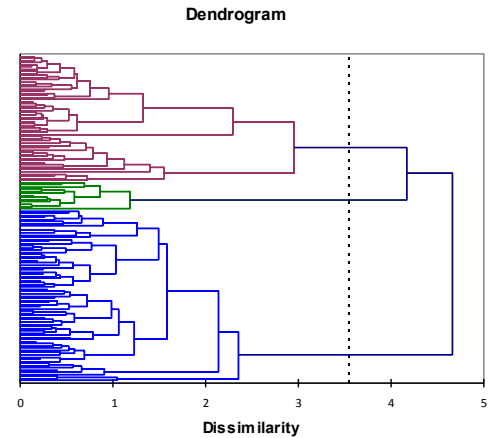


### Simple and Multiple Correspondence Analysis

- Problem, context of use and objectives
- Principle
- Frequency tables
- Chi-square decomposition
- Inertia and profiles
- Graphical tools
- Applications
- Hands-on exercises and group discussions

## Cluster Analysis

- Problem, context of use and objectives
- Steps involved in cluster analysis
- Notion of groups
- Data handling: Notion of distance and calculations
- Handling different types of data
- Variable selection and redundancy
- Standardization, weighing
- Handling missing values
- Methods and philosophies used to group objects
- Hierarchical methods: single linkage, complete linkage, average linkage, centroid
- Modeling methods: Ward and others
- Optimization methods: K-means and others
- Other methods: Fuzzy clustering
- Number of groups determination
- Graphical tools: Dendrograms, silhouette plots
- Statistical indicators
- Use of groups and cluster characterization
- Hands-on exercises and group discussions



## Discriminant Analysis

- Problem, context of use and objectives
- Steps involved in discriminant analysis
- Variable selection and redundancy
- Discriminant functions
- Assumptions
- Type of results
- Interpretation
- Hands-on exercises and group discussions

## Multivariate Technique for Relating two Datasets

- Canonical Analysis
- Workshop Conclusion
- Workshop : Questions and Answers (optional) : 14h00-16h30

**Observations (axes F1 and F2: 100.00 %)**

