Instrument Maintenance in the Fast Chromatography World

Brian G. Rohrback
Infometrix, Inc.

Gulf Coast Conference
October 2012

Poll of process users

1. Analytical failure prediction
2. Result validation
3. More process-specific information (timely, higher quality, more focused)
4. Simplification of procedures
5. Elimination of analytical discrepancies
6. Reduction in the lifecycle cost (cost of ownership)

Driven in part by dwindling manpower, skill sets and capabilities !

Another role for chemometrics

- With the increase in speed, we need to automate the assessment of the chromatographic data such that samples behaving normally are accepted, but any problem is noted whether it be
  - a raw material input deficiency,
  - a process problem, or
  - an instrument problem

Target knew a teen girl was pregnant before her father did

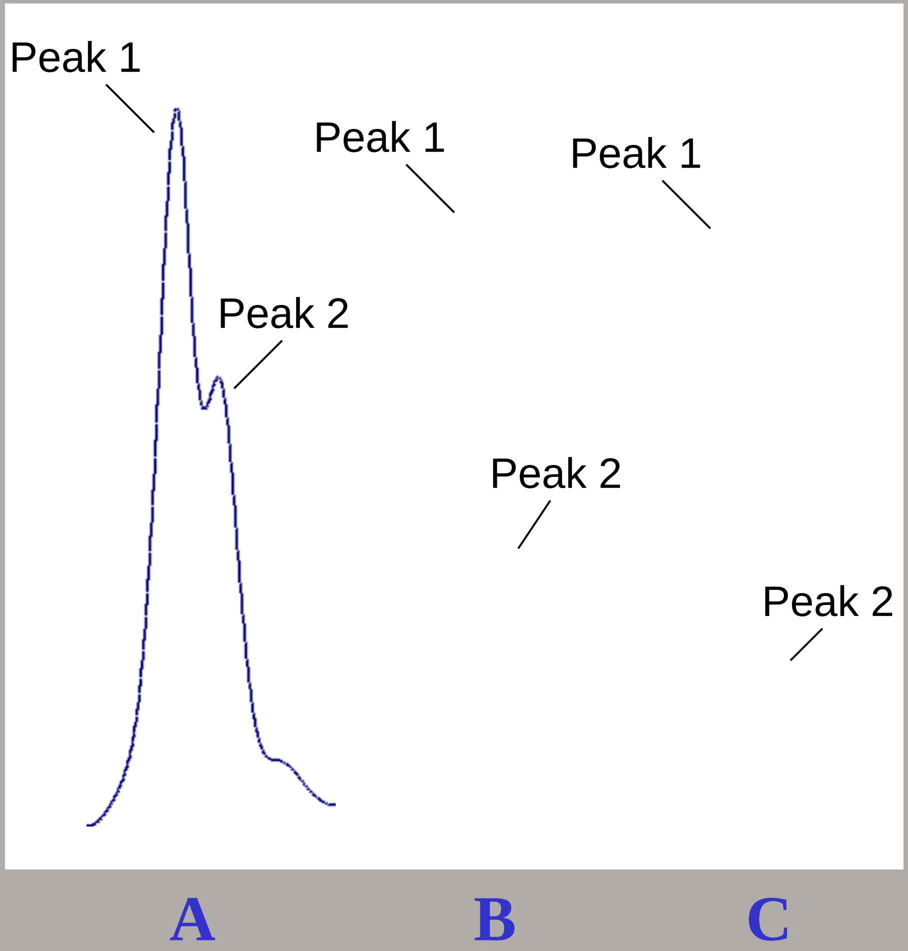
Every time you go shopping, you share intimate details about your consumption patterns with retailers.

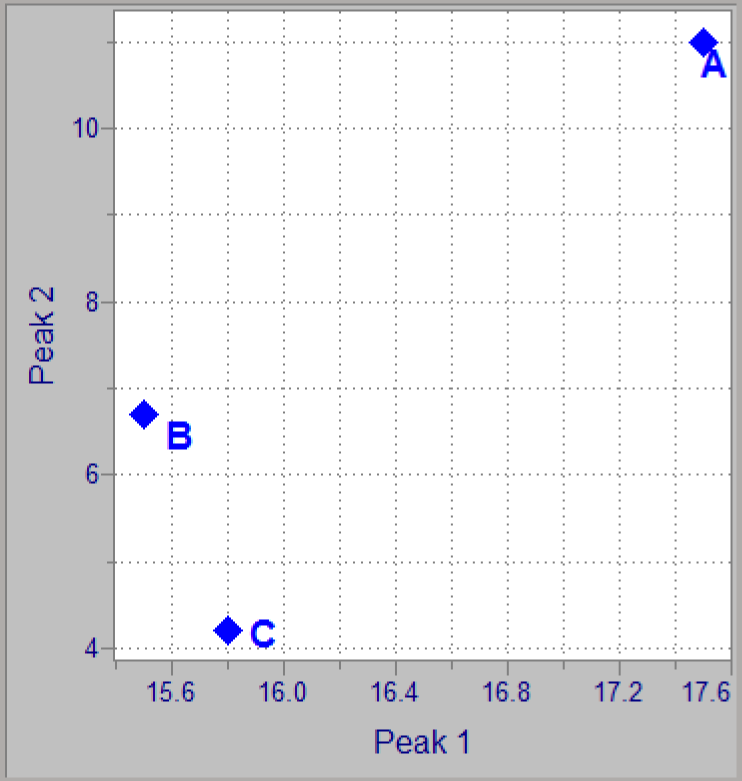
- A statistician looked at historical buying data for all the ladies who had signed up for Target baby registries in the past.
- Women on the baby registry were buying larger quantities of unscented lotion around the beginning of their second trimester.
- Sometime in the first 20 weeks, pregnant women loaded up on supplements like calcium, magnesium and zinc…

Target uses this not only to identify which customers are pregnant, but the timing of purchases can also estimate within weeks what the delivery date will be.

***How Companies Learn Your Secrets**, Charles Duhigg, NY Times, February 12, 2012*

3 points represent 3 chromatograms

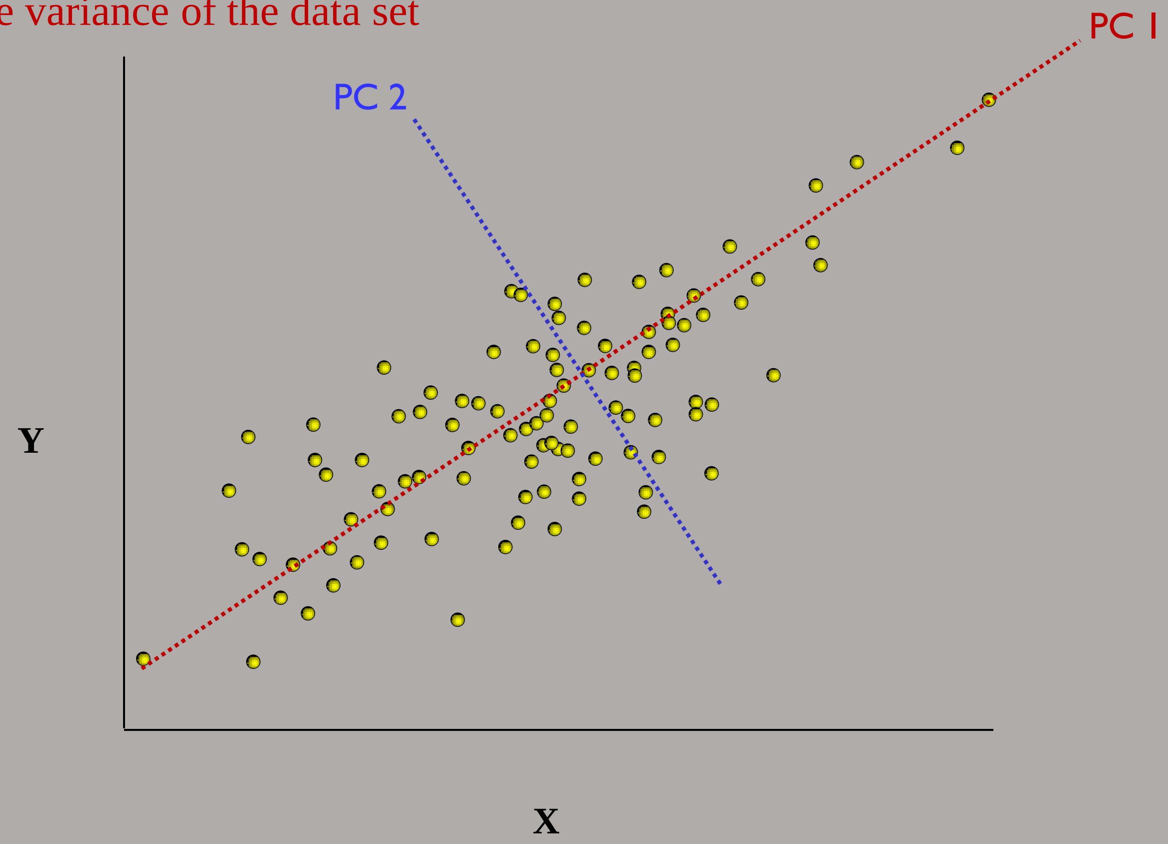




PCA method

First Principal Component (or Factor)

• Describes most of the variance of the data set



Second Principal Component (or Factor)

• Orthogonal to the First Principal Component

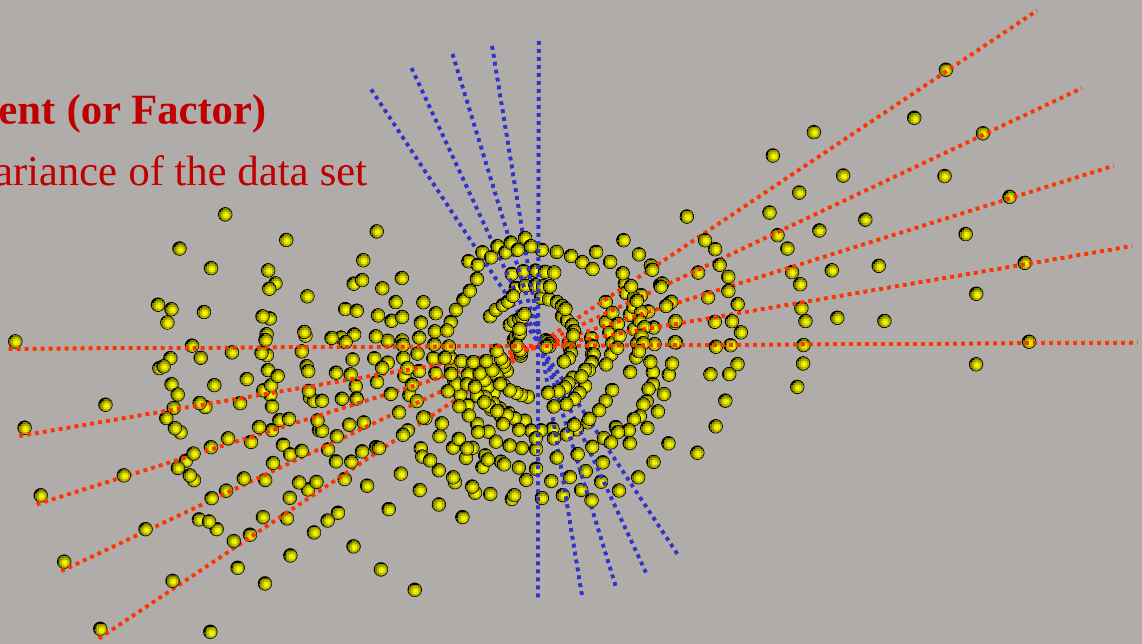
• Describes more of the variance of the data set not described by the First PC

PCA method

Rotation

First Principal Component (or Factor)

• Describes most of the variance of the data set

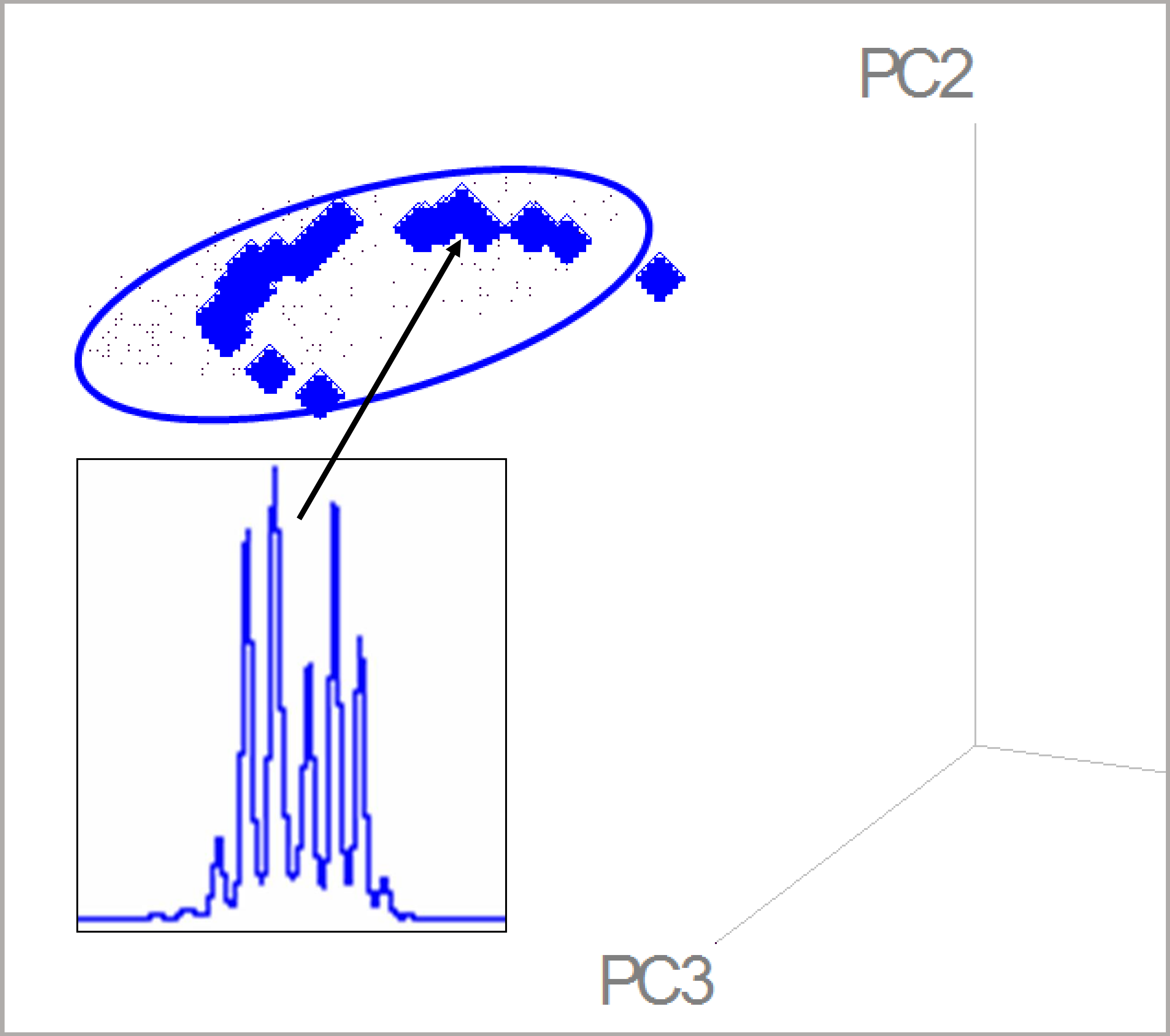


Second Principal Component (or Factor)

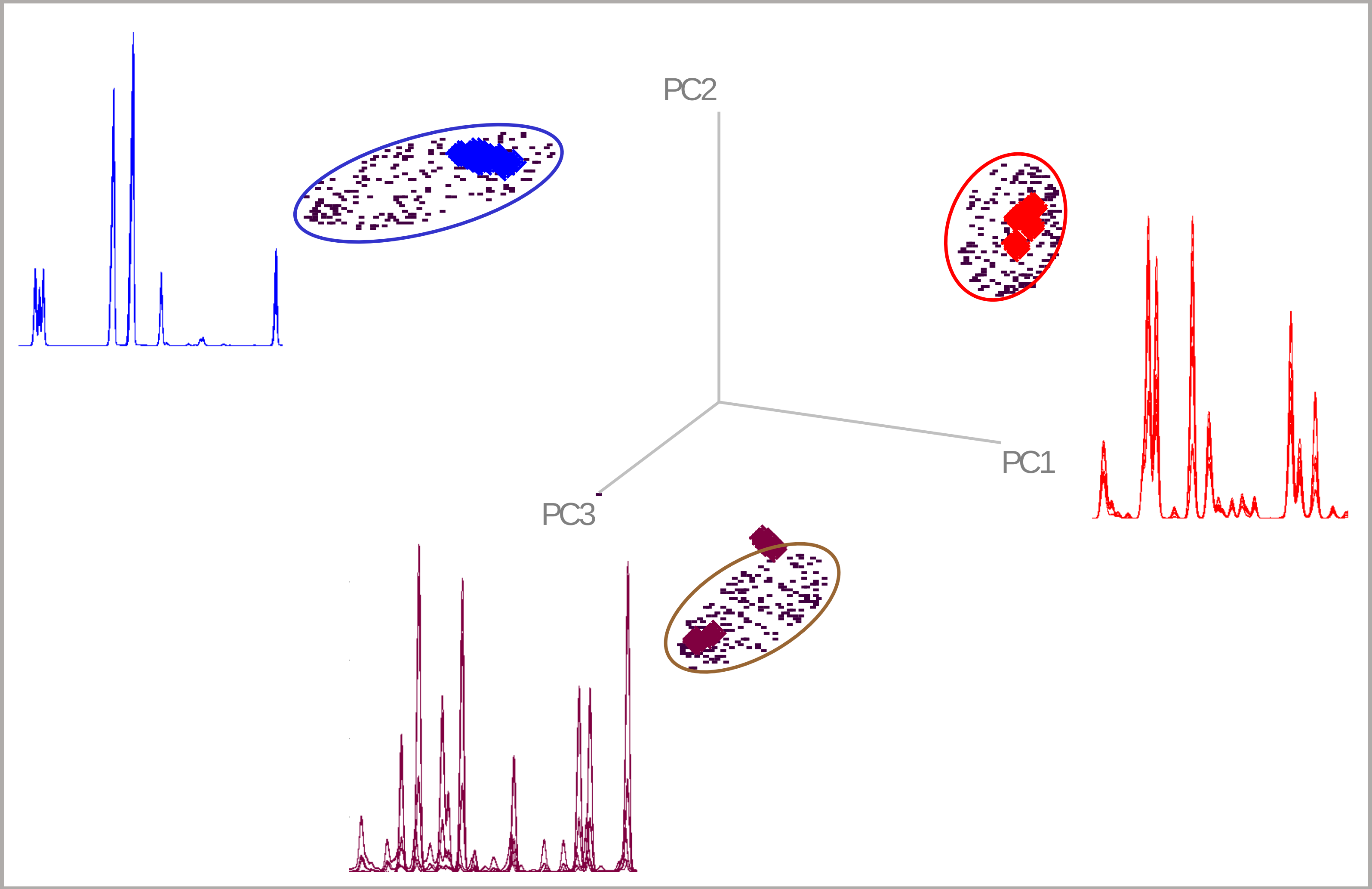
• Orthogonal to the First Principal Component

• Describes more of the variance of the data set not described by the First PC

A chromatogram is a point in PCA space



A basis for interpretation



Two Approaches

There are two ways of handling chromatographic data which can be done separately or in tandem:

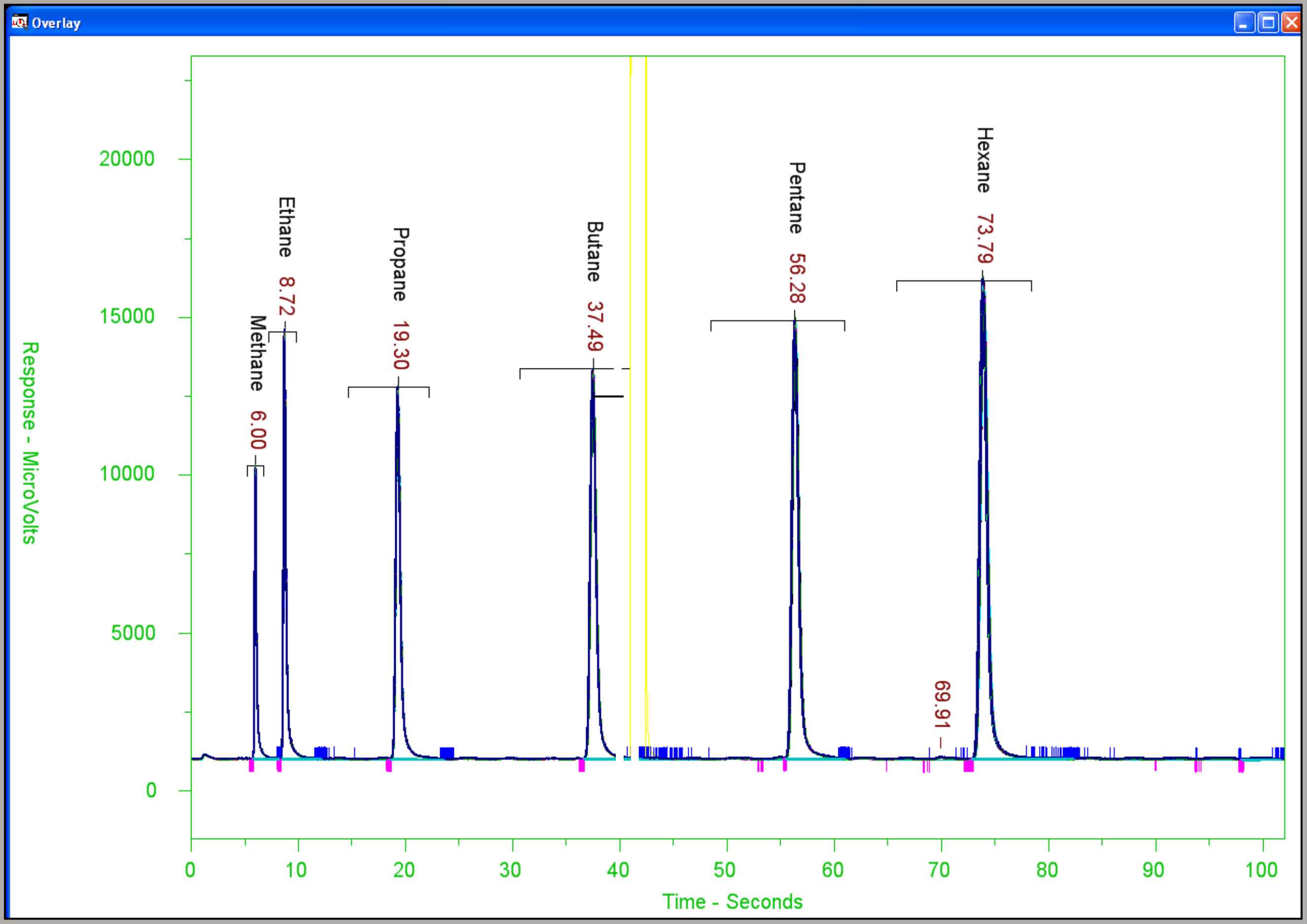
1. Peak Tables

   Peak tables are simple and small and contain concentration information about the primary components in the mixture. They miss unexpected peaks.

2. Raw Chromatograms

   Treating the chromatogram as if it were a spectrum means that both the expected and the unexpected are covered. This approach is sensitive to variations in retention time and forces the system to deal with 100 to 1000 times as much data.

Step 1: Consistent GC data



Step 2: Eliminate residual misalignment

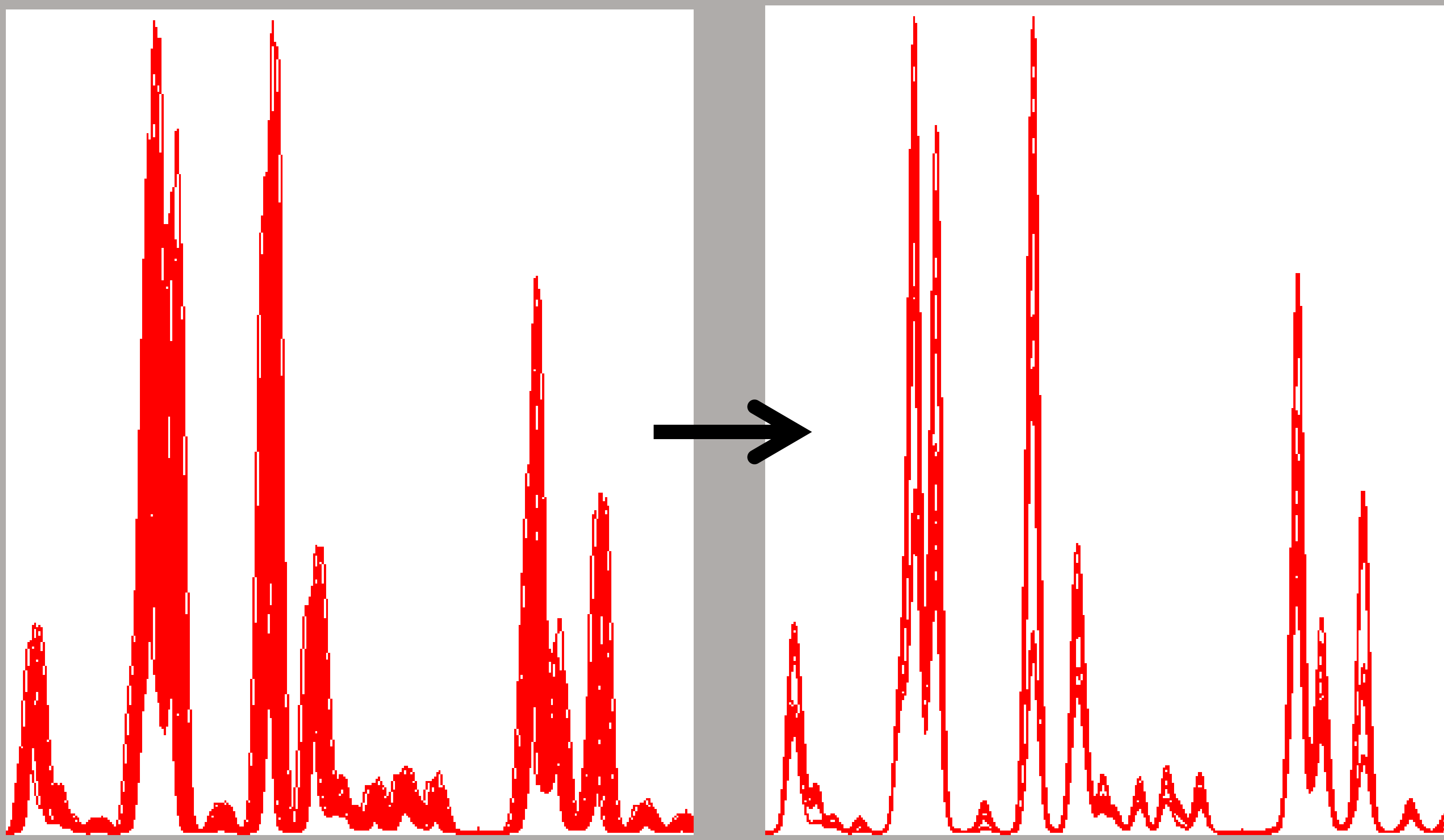


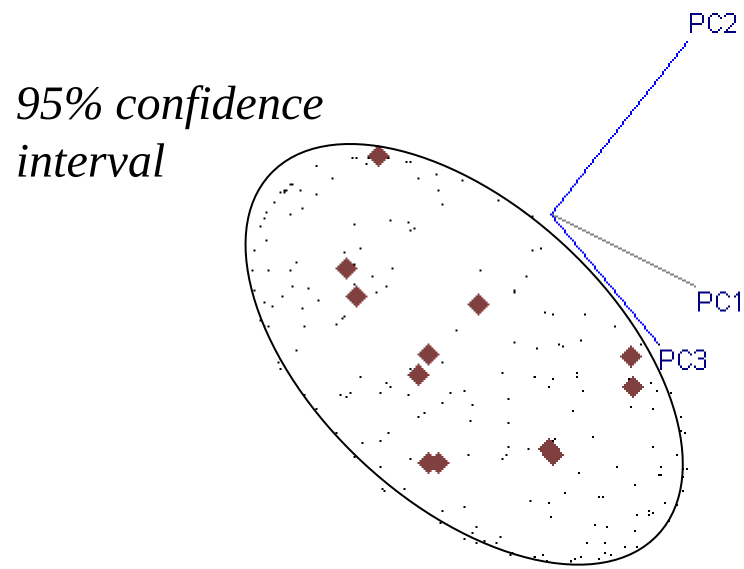
Table example: DHA summary report

Composite report					Winter gasoline		
Total by group type & carbon number							
(in volume percent)							
	n-Paraffins:	i-Paraffins:	Olefins:	Naphtenes:	Aromatics:	Oxygenates	Total:
C1:	0	0	0	0	0	0	0
C2:	0	0	0	0	0	0	0
C3:	0.424	0	0	0	0	0	0.424
C4:	11.108	4.513	1.1	0	0	0	16.721
C5:	0.427	2.867	0.928	0.029	0	0	4.251
C6:	0.197	2.112	0.467	0.348	0.198	0	3.323
C7:	1.054	14.459	17.395	0.411	10.5	0	43.819
C8:	0.404	3.686	0	0.827	10.715	0	15.631
C9:	0	0.938	0.052	1.667	6.04	0	8.698
C10:	0.141	1.915	0	0.208	2.24	0	4.505
C11:	0.024	0.951	0	0	0.125	0	1.101
C12:	0.035	0.068	0.025	0.026	0	0	0.155
C13:	0	0	0	0	0	0	0
C14:	0	0	0	0	0	0	0
Total:	13.816	31.509	19.966	3.517	29.818	0	98.626
Total C14+:	0						
Total unknowns:	1.374						
Grand total	100						

Infometrix

Step 3: Build a classification model

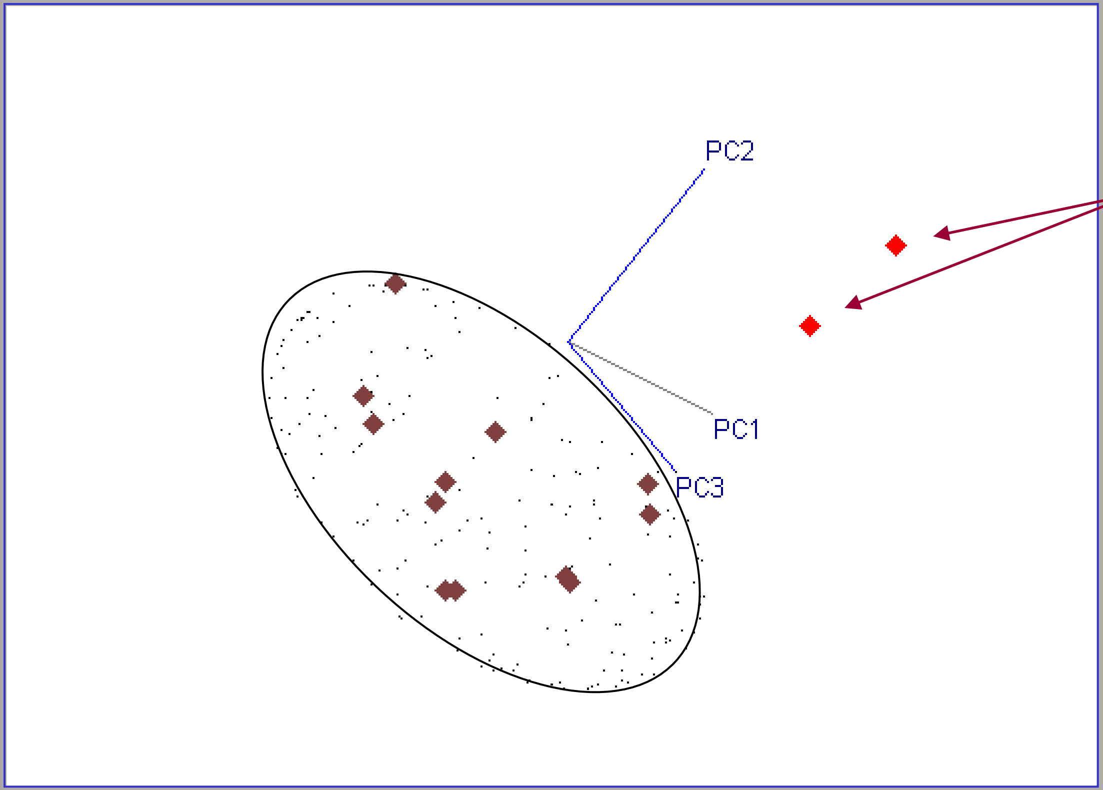
Composite report							
Total by group type & carbon number							
(in volume percent)							
	n-Paraffins:	i-Paraffins:	Olefins:	Naphtenes:	Aromatics:	Oxygenates:	Total:
C1:	0	0	0	0	0	0	0
C2:	0	0	0	0	0	0	0
C3:	0.424	0	0	0	0	0	0.424
C4:	11.108	4.513	1.1	0	0	0	16.721
C5:	0.427	2.867	0.928	0.029	0	0	4.251
C6:	0.197	2.112	0.467	0.348	0.198	0	3.323
C7:	1.054	14.459	17.395	0.411	10.5	0	43.819
C8:	0.404	3.686	0	0.827	10.715	0	15.631
C9:	0	0.938	0.052	1.667	6.04	0	8.698
C10:	0.141	1.915	0	0.208	2.24	0	4.505
C11:	0.024	0.951	0	0	0.125	0	1.101
C12:	0.035	0.068	0.025	0.026	0	0	0.155
C13:	0	0	0	0	0	0	0
C14:	0	0	0	0	0	0	0
Total:	13.816	31.509	19.966	3.517	29.818	0	98.626
Total C14+:	0						
Total unknowns:	1.374						
Grand total	100						



Composite report							
Total by group type & carbon number							
(in volume percent)							
	n-Paraffins:	i-Paraffins:	Olefins:	Naphtenes:	Aromatics:	Oxygenates:	Total:
C1:	0	0	0	0	0	0	0
C2:	0	0	0	0	0	0	0
C3:	0.424	0	0	0	0	0	0.424
C4:	11.108	4.513	1.1	0	0	0	16.721
C5:	0.427	2.867	0.928	0.029	0	0	4.251
C6:	0.197	2.112	0.467	0.348	0.198	0	3.323
C7:	1.054	14.459	17.395	0.411	10.5	0	43.819
C8:	0.404	3.686	0	0.827	10.715	0	15.631
C9:	0	0.938	0.052	1.667	6.04	0	8.698
C10:	0.141	1.915	0	0.208	2.24	0	4.505
C11:	0.024	0.951	0	0	0.125	0	1.101
C12:	0.035	0.068	0.025	0.026	0	0	0.155
C13:	0	0	0	0	0	0	0
C14:	0	0	0	0	0	0	0
Total:	13.816	31.509	19.966	3.517	29.818	0	98.626
Total C14+:	0						
Total unknowns:	1.374						
Grand total	100						

Step 4: Identify cause for outlier groupings and trends

Composite report							
Total by group type & carbon number							
(in volume percent)							
	n-Paraffins:	i-Paraffins:	Olefins:	Naphtenes:	Aromatics:	Oxygenates	Total:
C1:	0	0	0	0	0	0	0
C2:	0	0	0	0	0	0	0
C3:	0.424	0	0	0	0	0	0.424
C4:	11.108	4.513	1.1	0	0	0	16.721
C5:	0.427	2.867	0.928	0.029	0	0	4.251
C6:	0.197	2.112	0.467	0.348	0.198	0	3.323
C7:	1.054	14.459	17.395	0.411	10.5	0	43.819
C8:	0.404	3.686	0	0.827	10.715	0	15.631
C9:	0	0.938	0.052	1.667	6.04	0	8.698
C10:	0.141	1.915	0	0.208	2.24	0	4.505
C11:	0.024	0.951	0	0	0.125	0	1.101
C12:	0.035	0.068	0.025	0.026	0	0	0.155
C13:	0	0	0	0	0	0	0
C14:	0	0	0	0	0	0	0
Total:	13.816	31.509	19.966	3.517	29.818	0	98.626
Total C14+:	0						
Total unknowns:	1.374						
Grand total	100						

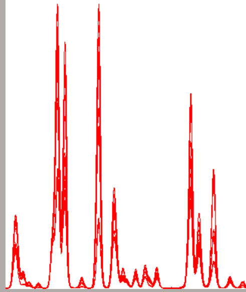


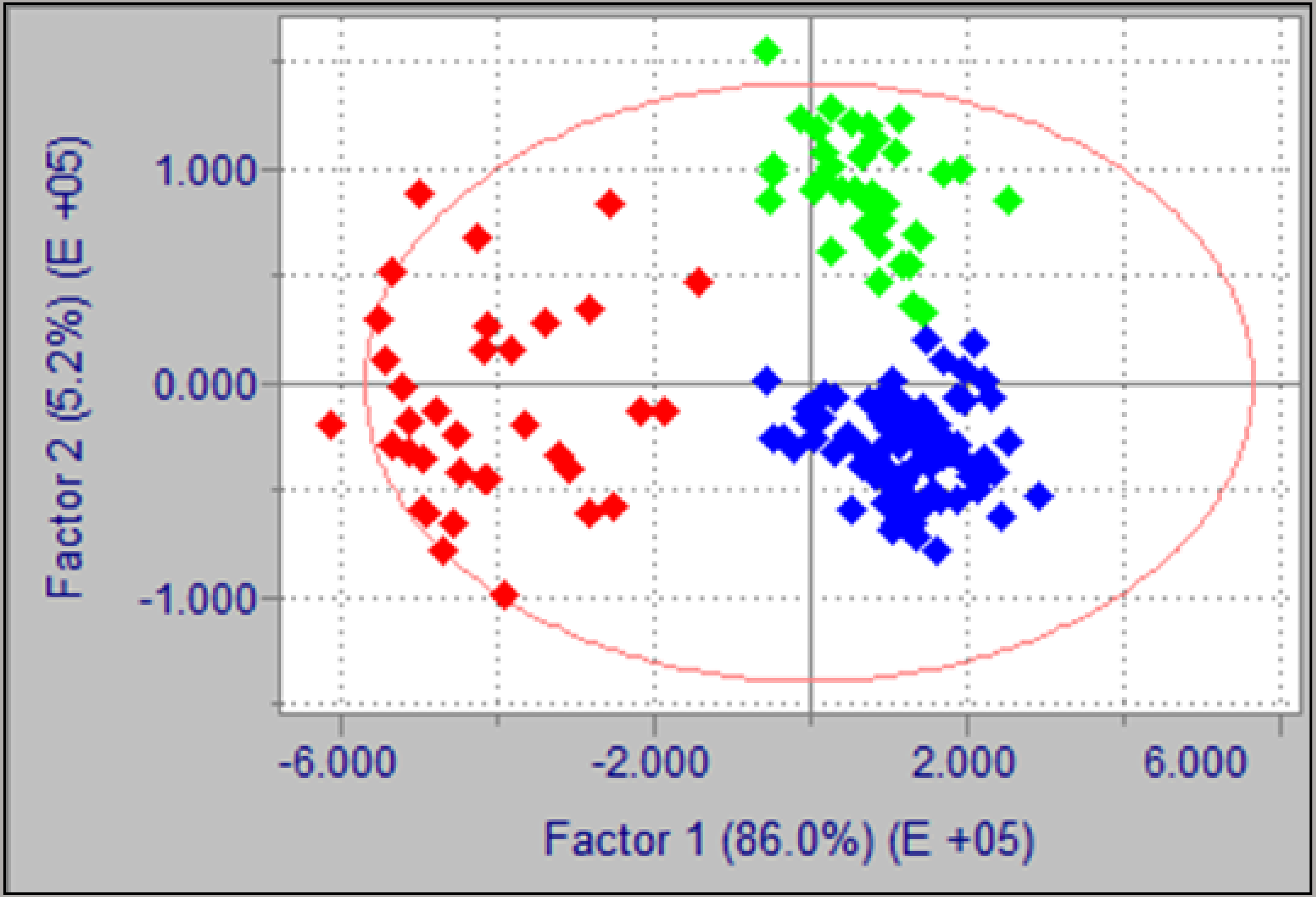
Outliers, because of:

- Instrument problem?
- Process upset?

Infometrix

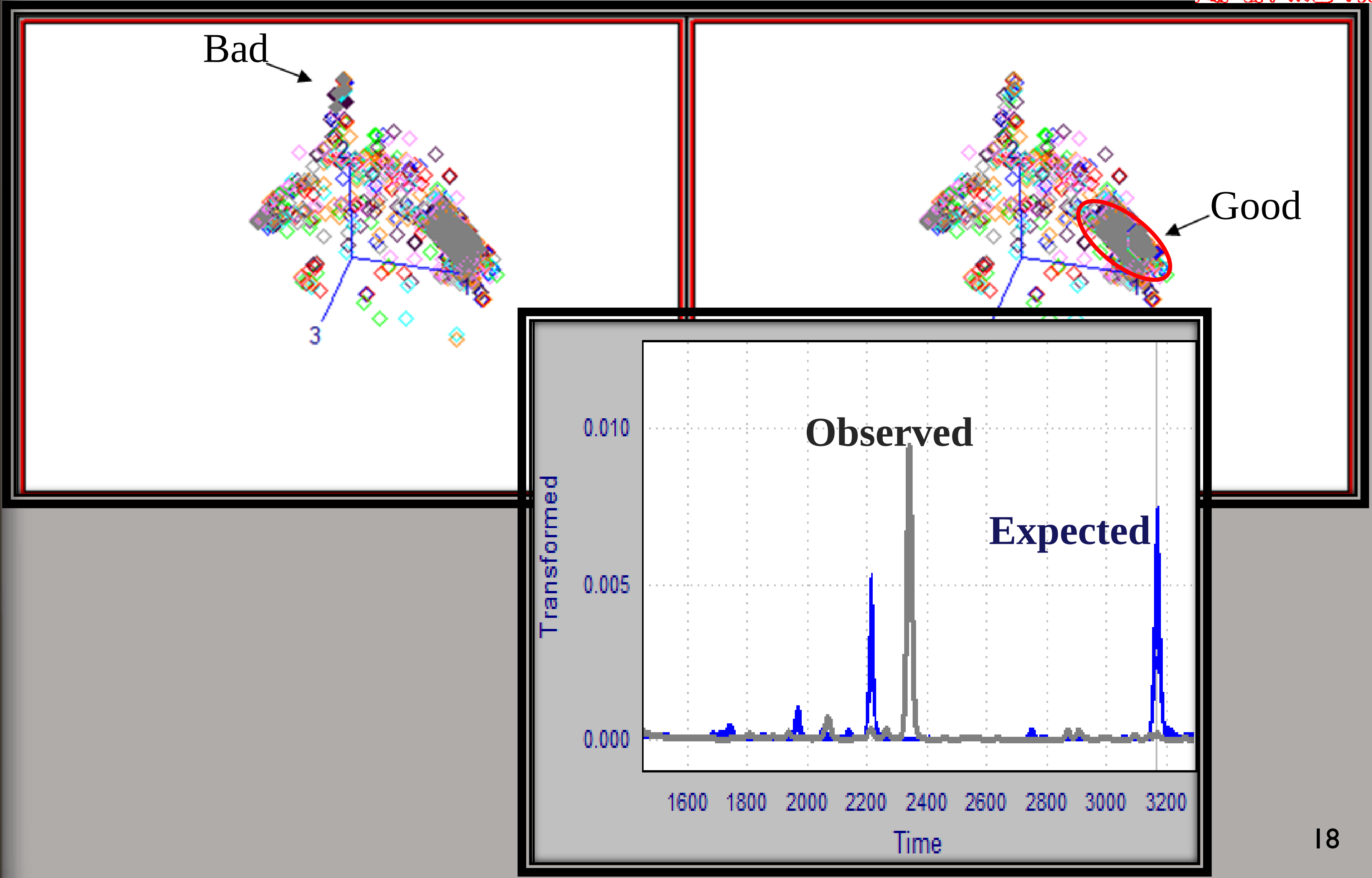
Step 3: Build a classification model using aligned chromatograms





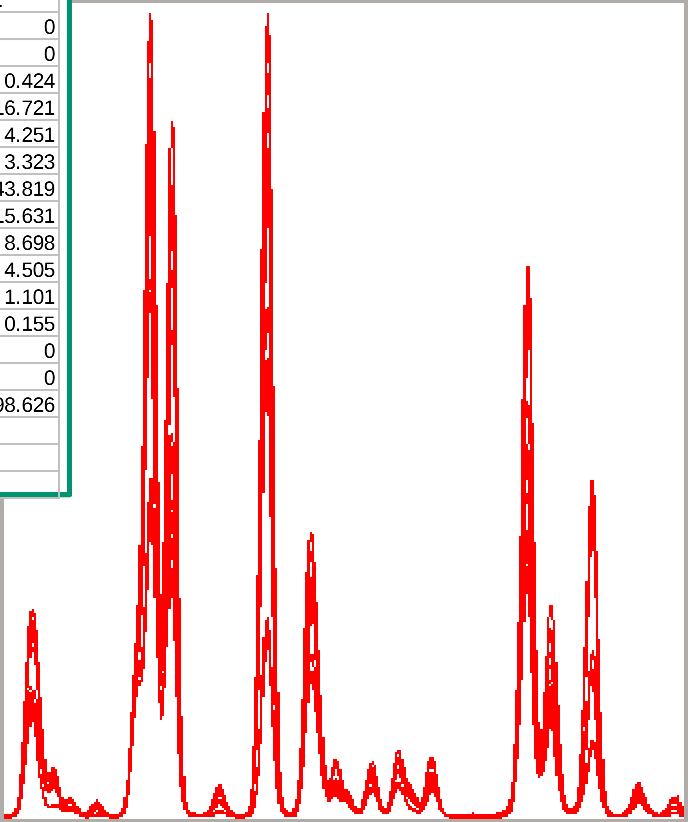
We selected 160 representative good-quality chromatograms to make a model from the 3600 chromatograms supplied.

Step 4: Identify cause for outlier groupings and trends



Good to Go

Composite report							
Total by group type & carbon number							
(in volume percent)							
	n-Paraffins:	i-Paraffins:	Olefins:	Naphtenes:	Aromatics:	Oxygenates	Total:
C1:	0	0	0	0	0	0	0
C2:	0	0	0	0	0	0	0
C3:	0.424	0	0	0	0	0	0.424
C4:	11.108	4.513	1.1	0	0	0	16.721
C5:	0.427	2.867	0.928	0.029	0	0	4.251
C6:	0.197	2.112	0.467	0.348	0.198	0	3.323
C7:	1.054	14.459	17.395	0.411	10.5	0	43.819
C8:	0.404	3.686	0	0.827	10.715	0	15.631
C9:	0	0.938	0.052	1.667	6.04	0	8.698
C10:	0.141	1.915	0	0.208	2.24	0	4.505
C11:	0.024	0.951	0	0	0.125	0	1.101
C12:	0.035	0.068	0.025	0.026	0	0	0.155
C13:	0	0	0	0	0	0	0
C14:	0	0	0	0	0	0	0
Total:	13.816	31.509	19.966	3.517	29.818	0	98.626
Total C14+:	0						
Total unknowns:	1.374						
Grand total	100						



Infometrix

Chemometrics for instrumentation: the value proposition

Anything you can do to improve precision of the multivariate measurements collected by the instrument will allow you to tighten the control – essentially for free.

We use the signal processing aspect of chemometrics to reduce instrument-derived variability

- ***Within an instrument (e.g., noise reduction)***
- ***Between instruments (i.e., transfer of calibration)***

This creates the ability to construct an application-specific, objective evaluation system