Predictive Analytics for New Product Forecasting

More *Analytics* and less *Judgment*: a Case Study

Dr. Sven F. Crone
Dr. Nikolaos Kourentzes

Lancaster Centre for Forecasting
www.forecasting-centre.com

New Product Forecasting = Gambling?

Both?

Statistics?

e.g. Norton-Bass

Classification
Clustering
Social-media analysis
…

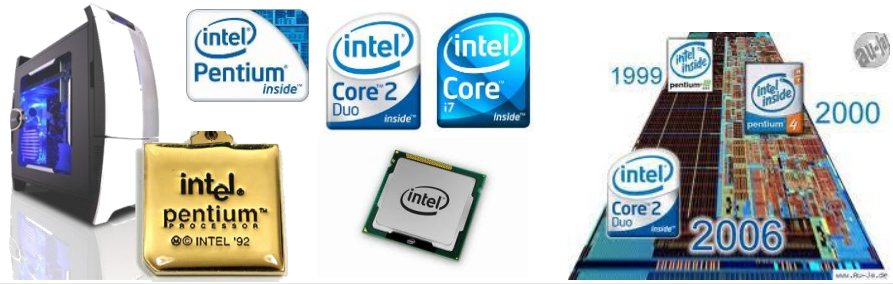
Judgment?

e.g. Delphi
Szenarios
Market Tests
Prediction Markets

→ Statistics vs. Judgment in New Product Forecasting?

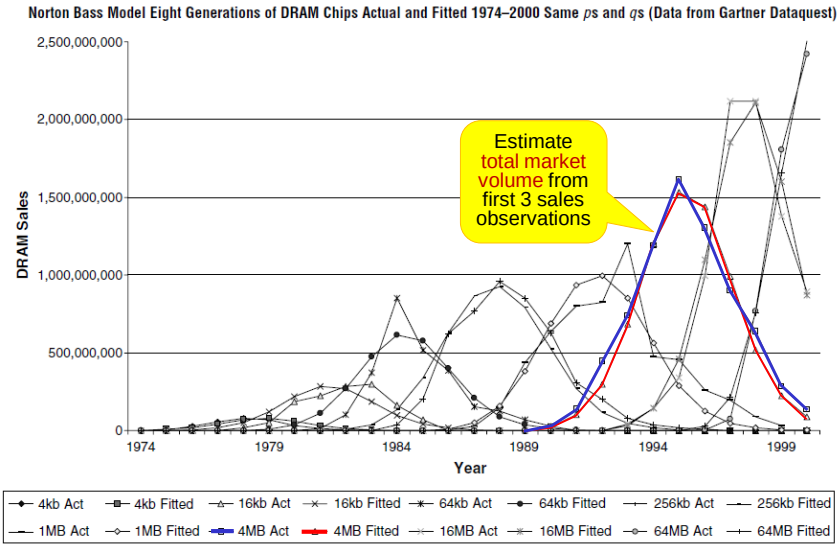
The Challenge

- How to forecast high-tech products (CPU,RAM) at Intel?
  - Product sourcing & production lead times are longer than sales period → every product is "new"
  - Often requires building new production facilities (or even plants) → high costs for under & overforecasting
  - 100s to 1,000s of new products per season
  - 2-4 innovation cycles (selling seasons) per year → Requires semi-automated & standardized process
  - How to they do New product forecasting?



Norton-Bass Model

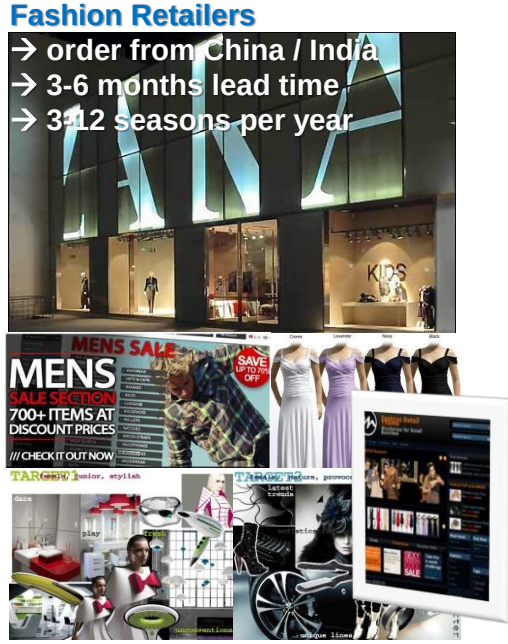
Sales of products with continuous repeat purchasing?



Common Problem

Fashion Retailers

→ order from China / India
→ 3-6 months lead time
→ 3-12 seasons per year



Catalogue Retailers

→ order from China / India
→ 3-6 months lead time
→ 4 catalogues per year

Sourcing lead times are longer than sales period (for many products)
→ many products are "new"

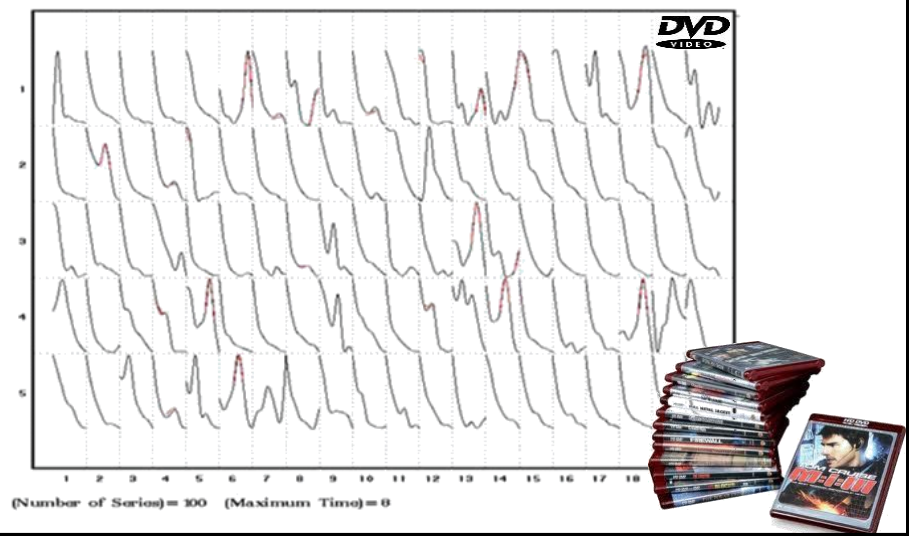
Common Problem

- CDs, DVDs, BlueRays, Video Games …

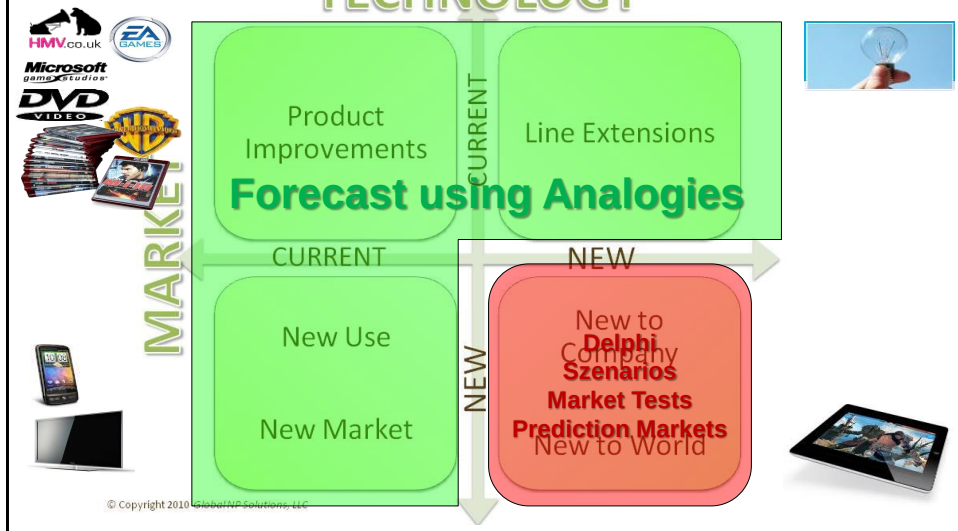
→**Product launch & intial sales before, during & after go-live (i.e. pipeline-filling) are most important**

Common Problem

… check out the scaled thumbnails of the first 8 weeks of sales time series plots for 100 new DVD releases



Few things are completely new!



→How to find similar products from history (aka Analogies) automatically?

Finding Analogies: you try it!

aka statistical technique called „Clustering“

Comparing Comparable Things

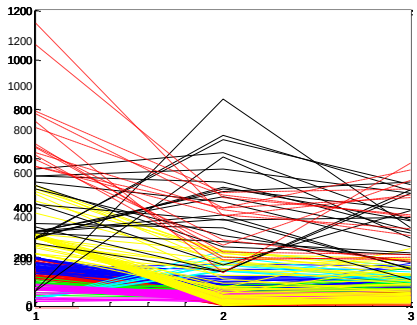
Cluster 1: Apples

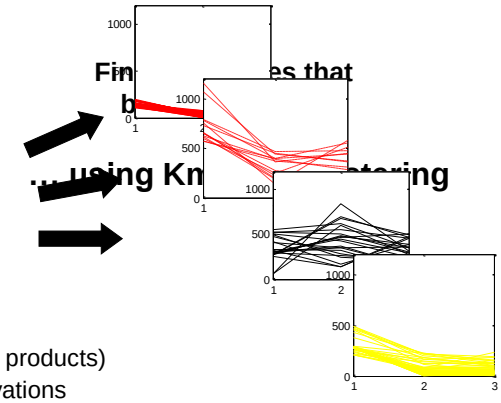
Cluster 2: Oranges

Now find:

Analogies in Time Series?

Hard to extract useful information! → No additional information





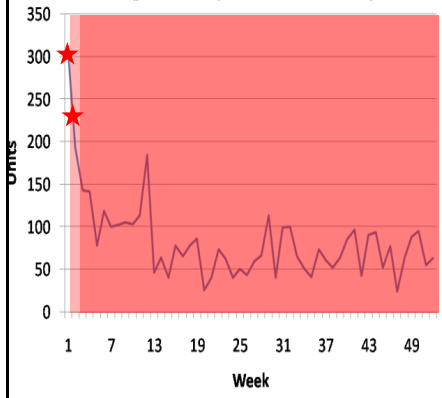
Focus on the first three months (new products)
Time series with 27 months of observations

→ Finds clusters of typical initial sales behavior

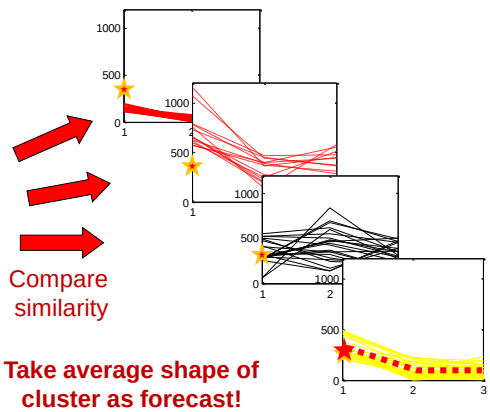
Analogies in Time Series?

New product to be launched

Multiple observations …
(plus 10s of product attributes etc.)



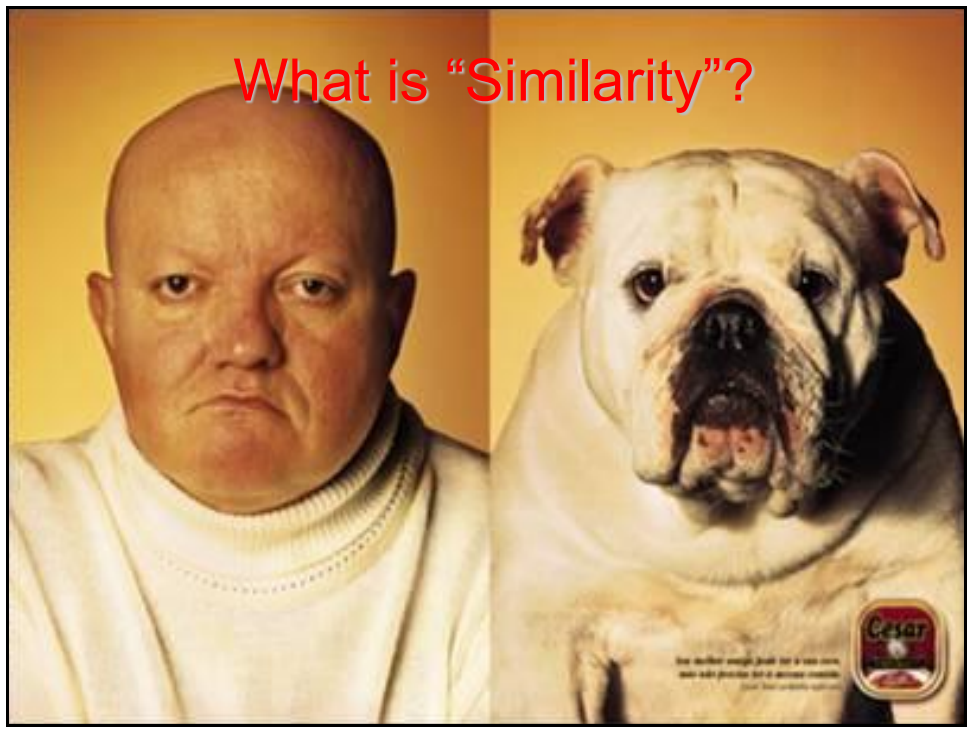
→ Match to *most similar* cluster

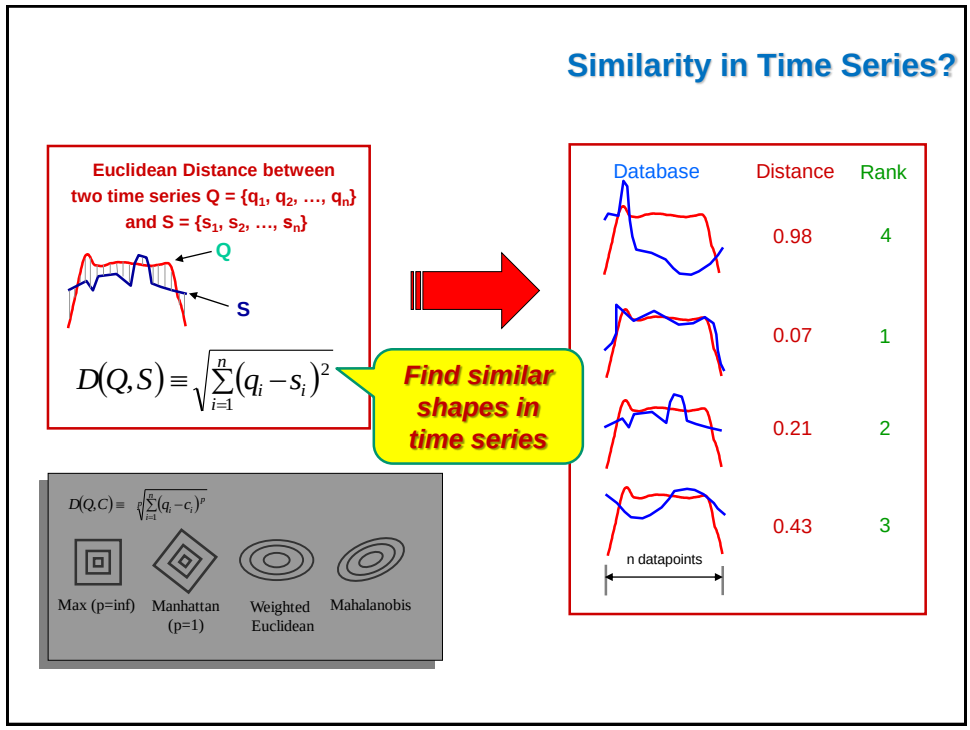


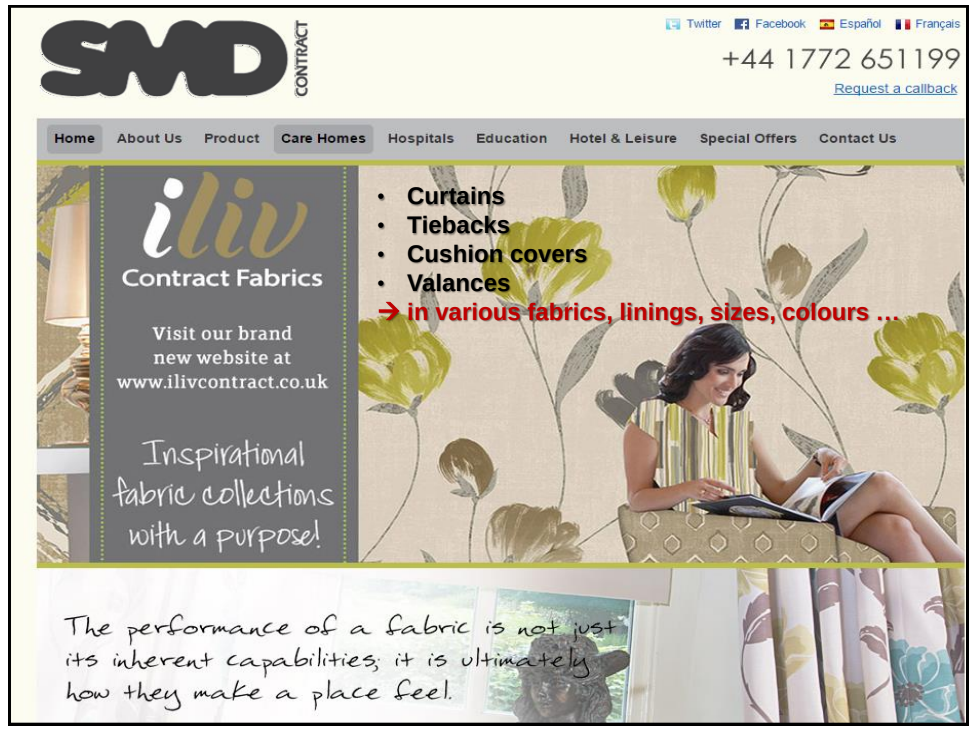
Compare similarity

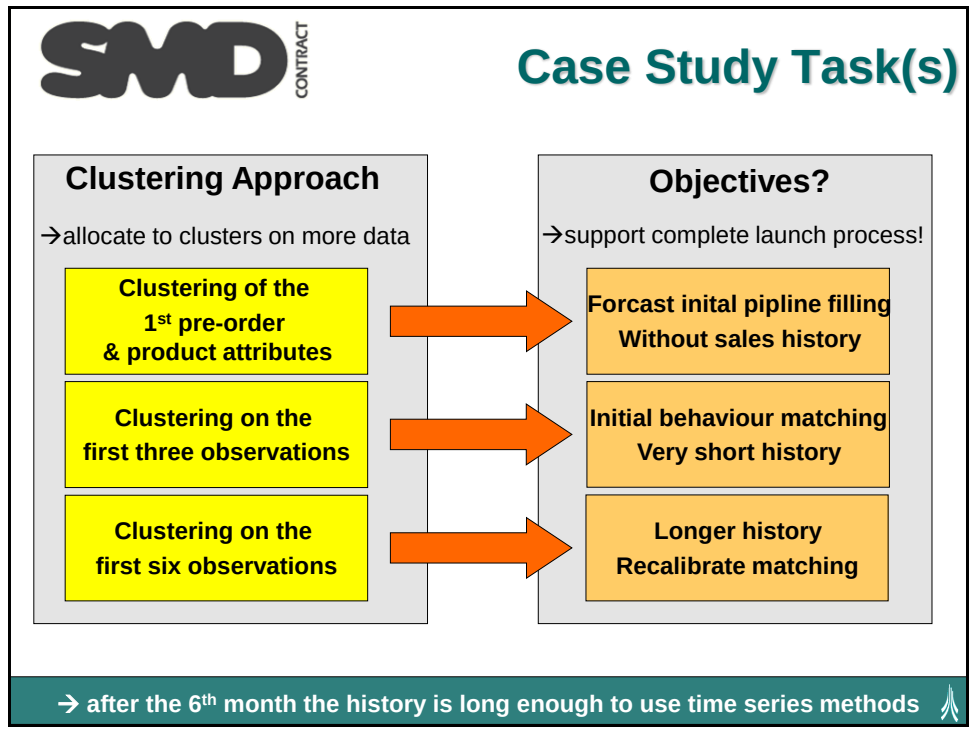
Take average shape of cluster as forecast!

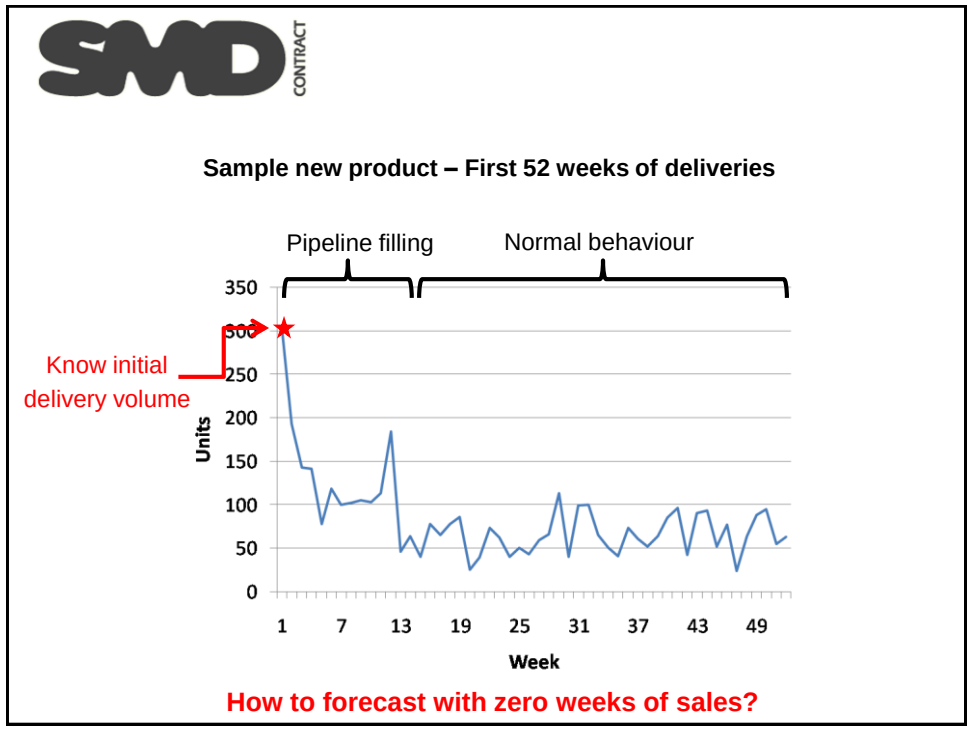
→ Need to be careful in specifying ***SIMILIARITY*** !

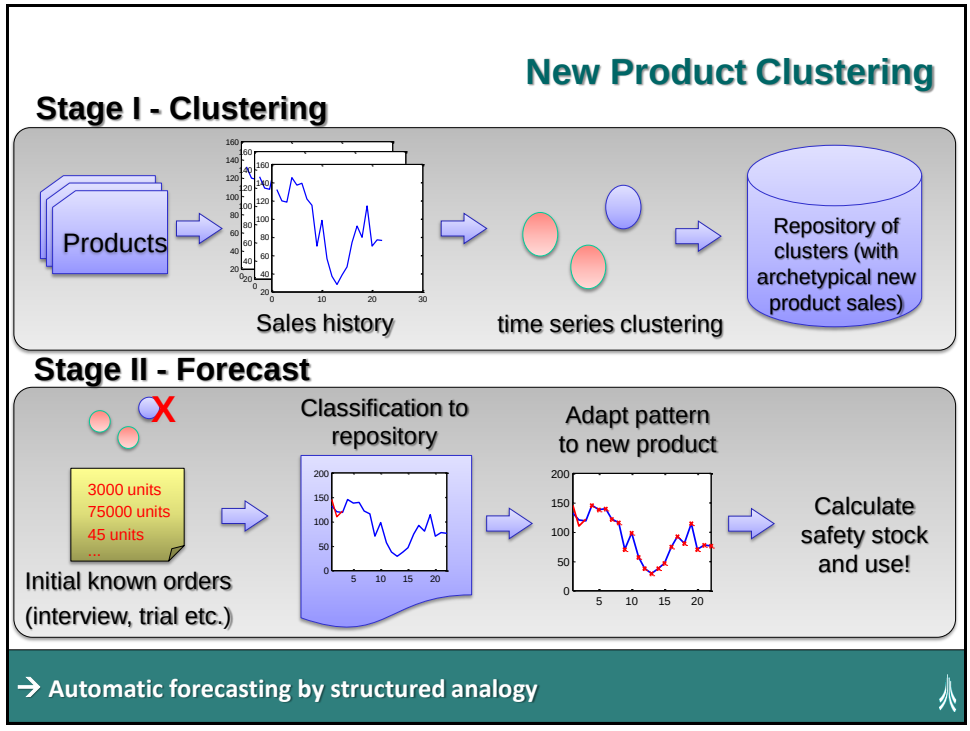












Clustering Results
3 first observations

I. **262 products used** with history up to 27 months

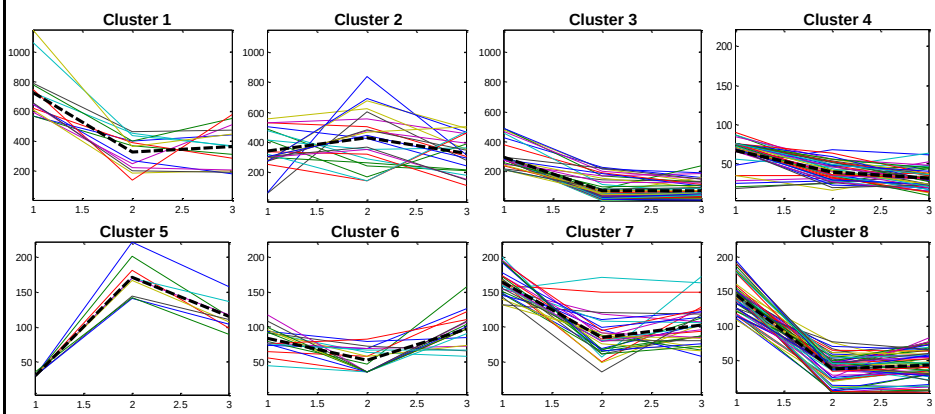
II. Each product → **Only first three months** used

III. **8 clusters** → Optimal number

IV. Each cluster → Average behaviour found

V. Average behaviour → New products will be matched with

Clustering Results
3 first observations



	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8
CC	64.29%	59.09%	3.57%					
TBKS	35.71%	40.91%	16.07%					
UL			80.36%	13.04%	100.00%	16.67%	76.67%	46.27%
LD				82.61%		83.33%	16.67%	47.76%
VAL				4.35%			6.67%	5.97%

Business intelligence can eliminate/correct clusters → **Cluster 5 case**

Forecasting

I. **Using the estimation for the first month value and the product type find the correct cluster using the three first observations clustering**

e.g. First value = 150, Unlined curtain → Cluster 7!

Cluster	Population		Shape			Shape in % change		
			1st Value	2nd Value	3rd Value	1st Value	2nd Value	3rd Value
Cluster 1	14	5.34%	726.14	324.14	365.21	-	-55.36%	12.67%
Cluster 2	22	8.40%	343.14	427.95	324.27	-	24.72%	-24.23%
Cluster 3	56	21.37%	296.20	66.43	66.02	-	-77.57%	-0.62%
Cluster 4	46	17.56%	66.74	39.50	30.83	-	-40.81%	-21.96%
Cluster 5	9	3.44%	30.67	170.89	115.44	-	457.25%	-32.44%
Cluster 6	18	6.87%	83.17	52.28	96.39	-	-37.14%	84.38%
Cluster 7	30	11.45%	164.13	84.97	102.33	-	-48.23%	20.44%
Cluster 8	67	25.57%	144.21	37.34	42.60	-	-74.10%	14.07%

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8
CC	64.29%	59.09%	3.57%					
TBKS	35.71%	40.91%	16.07%					
UL			80.36%	13.04%	100.00%	16.67%	76.67%	46.27%
LD				82.61%		83.33%	16.67%	47.76%
VAL				4.35%			6.67%	5.97%

Forecasting

I. **Using the estimation for the first month value and the product type find the correct cluster using the three first observations clustering**

II. **Use the % change of the average shape of the cluster to form the 2nd and 3rd value**

Cluster	Population		Shape			Shape in % change		
			1st Value	2nd Value	3rd Value	1st Value	2nd Value	3rd Value
Cluster 1	14	5.34%	726.14	324.14	365.21	-	-55.36%	12.67%
Cluster 2	22	8.40%	343.14	427.95	324.27	-	24.72%	-24.23%
Cluster 3	56	21.37%	296.20	66.43	66.02	-	-77.57%	-0.62%
Cluster 4	46	17.56%	66.74	39.50	30.83	-	-40.81%	-21.96%
Cluster 5	9	3.44%	30.67	170.89	115.44	-	457.25%	-32.44%
Cluster 6	18	6.87%	83.17	52.28	96.39	-	-37.14%	84.38%
Cluster 7	30	11.45%	164.13	84.97	102.33	-	-48.23%	20.44%
Cluster 8	67	25.57%	144.21	37.34	42.60	-	-74.10%	14.07%

Second Value (Month): → Third Value **(Month):**

150 * (1-0,4823) = 77.65 → 77.65 * (1+0,2044) = 93.52

Forecasting

I. Using the estimation for the first month value and the product type find the correct cluster using the three first observations clustering.

II. Use the % change of the average shape of the cluster to form the 2nd and 3rd value.

III. For the 4th-6th month forecasts use the clustering of the first 6 months to find the product membership (similar procedure!).

IV. Update the forecasted values with real whenever they are available. If possible re-evaluate membership when new values are available.

V. To forecast after the sixth value/month use a statistical model. The product history is now long enough!

New Product Forecasting

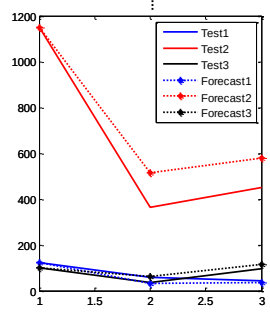
Experimental Evaluation

How accurate is it?

Evaluate on time series of the given data

Proposed Approach

	Mean Absolute Error	Mean Absolute Percentage
Test1	11.3	21%
Test2	92.3	23%
Test3	15.0	32%



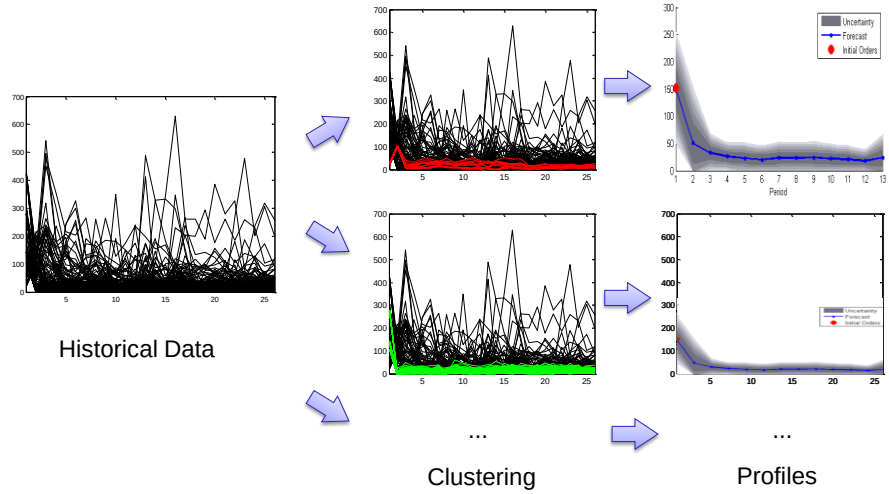
Statistical Forecasting

	Mean Absolute Error	Mean Absolute Percentage
Test1	47.0	95%
Test2	495.0	123%
Test3	22.0	62%

→ Time series clustering always outperforms statistical approach
→ Time series clustering outperformed human demand planner

New Product Clustering

Allows us to measure safety stock!



Historical Data

...

Clustering

...

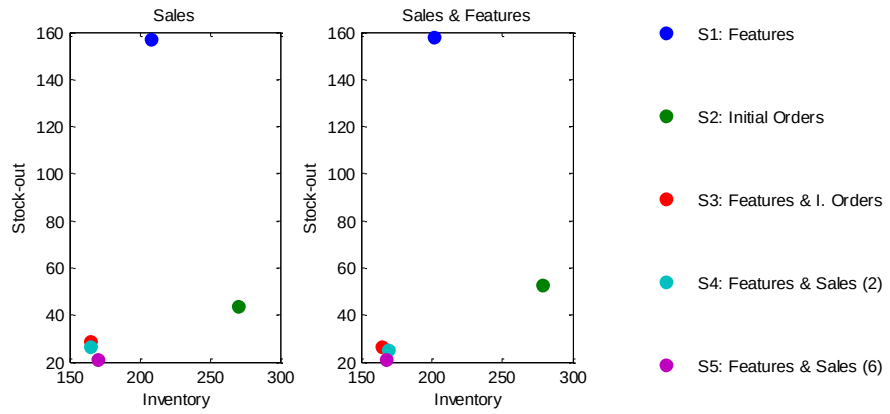
Profiles

Empirical Evaluation

Inventory Performance

Inventory Performance Results

Clustering on



Clustering Results
A simple Forecasting Tool

→ Outcome: simple MS EXCEL application that can be automated

Take aways

Looking for companies for a free-of-charge pilot study (to present at IBF?)

- Many industries specialize on New Product Forecasting → opportunities to learn (steal fire)
- Clustering creates insightful groups of similar products → helpful to planners to improve judgment
- Allows large-scale automation of new product forecasting
- Predictive Analytics provides new solutions to forecasting challenges

… room for improvement!

[LCF ISF09, SAS New Product Presentation ISF 2008]