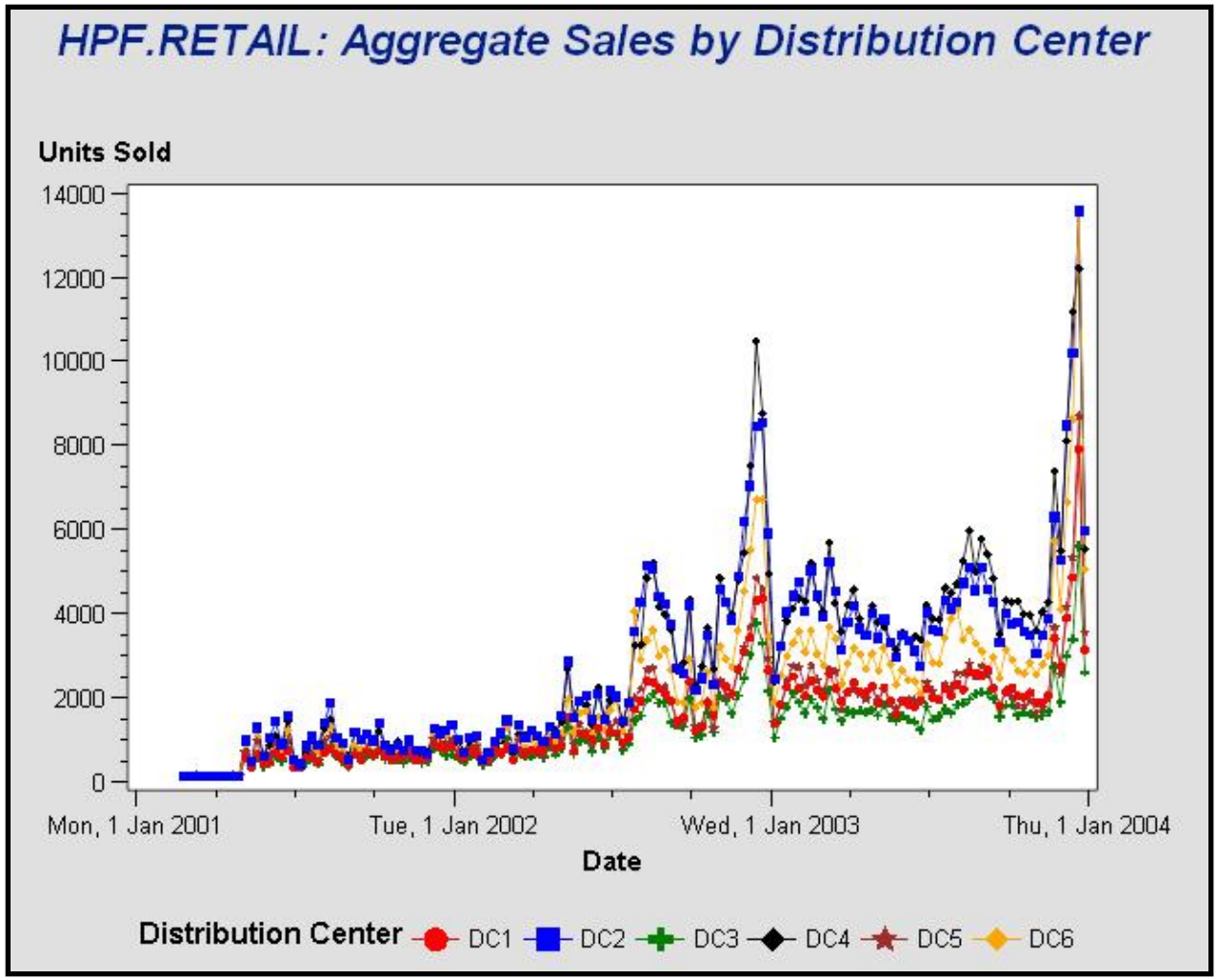
HIERARCHICAL AND GROUPED TIME SERIES

Dr. Aric LaBarr

Institute for Advanced Analytics

LARGE SCALE FORECASTING

Motivation


HPF.RETAIL: Aggregate Sales by Distribution Center
Units Sold
14000
12000
10000
8000
6000
4000
2000
0
Mon, 1 Jan 2001
Tue, 1 Jan 2002
Wed, 1 Jan 2003
Thu, 1 Jan 2004
Date
Distribution Center
DC1
DC2
DC3
DC4
DC5
DC6

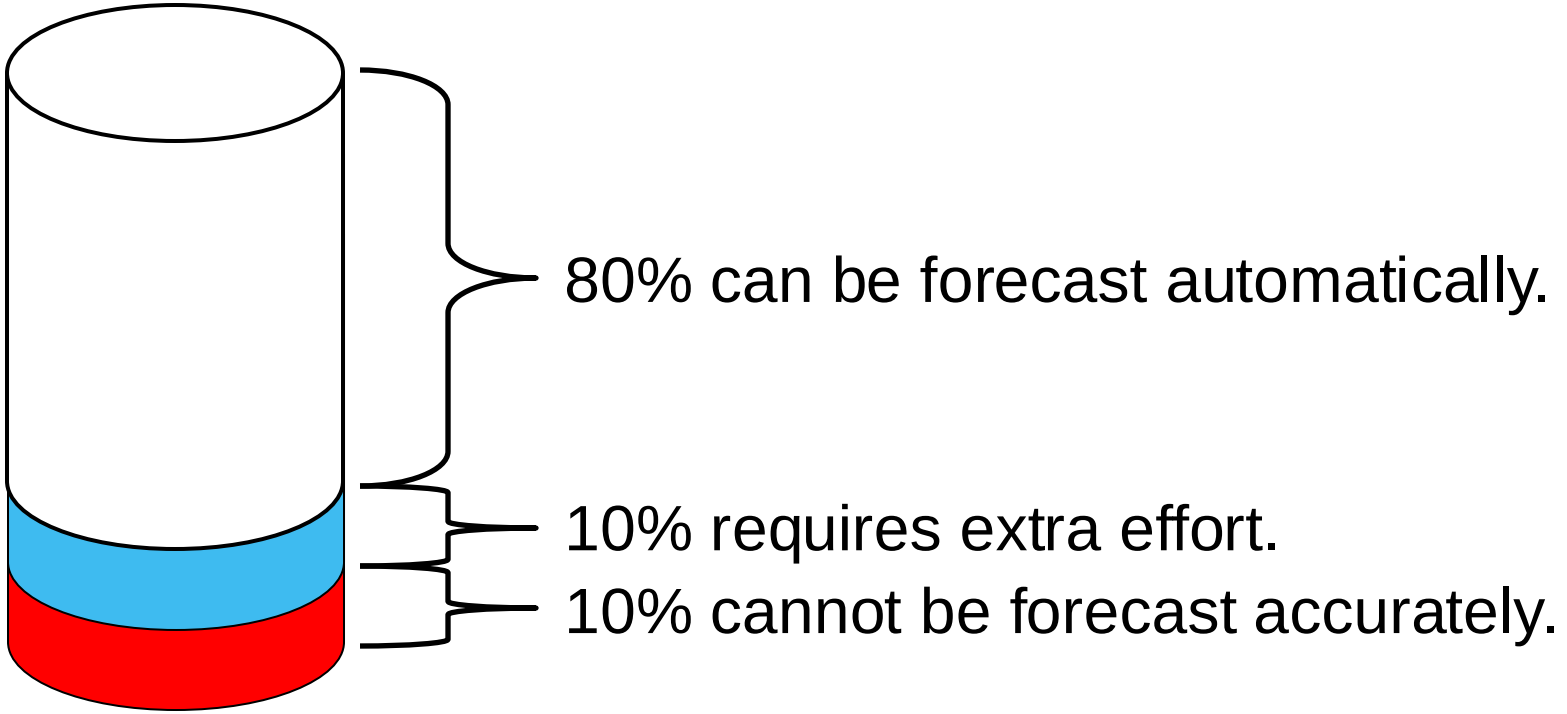
Forecasting a Single Time Series

- Skilled analysts can forecast an individual time series by one of the following:
  - Applying good judgement based on knowledge and experience
  - Use proven time series analysis techniques
  - Combination of both!

Large-Scale Forecasting

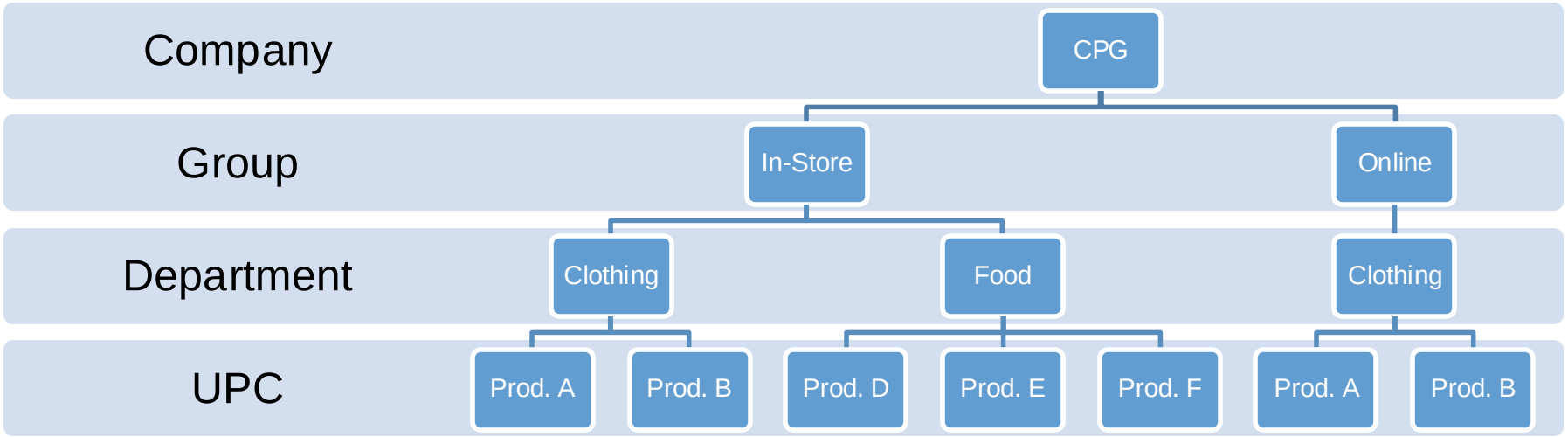
- Modern businesses require efficient, reliable forecasts for many series, not just one.
- Not sufficient resources to apply the same individualized approaches to all the series that need to be forecasted.
- Series might have hierarchically arranged elements and require reconciliation of forecasts at different levels.

Large-Scale Forecasting



HIERARCHICAL FORECASTING

Data Hierarchies



Automatic Model Building

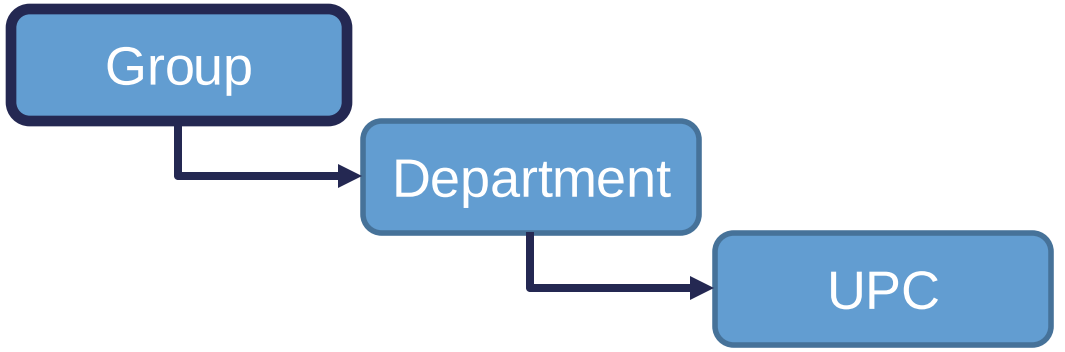
- Each model inside a certain hierarchy could easily be different.
- Most software can do this “automatically”:
  - SAS Forecast Studio uses the model that “fits the data the best.”
    - Builds models based on MAPE using automatic selection techniques.
  - R fable package
    - Can define structures of hierarchies across columns of data.
    - Builds all models contained in ‘model’ functionality

Model Reconciliation

- If we were to model each series in each hierarchy, the statistical forecasts typically wouldn't add up to between hierarchies as we would want.
- Example:
  - If we sell 20 products in a region, we want our region forecast to be the sum of the 20 product forecasts in that region.
- **Reconciliation** is the process of making the statistical forecasts add up for each time interval in the data.

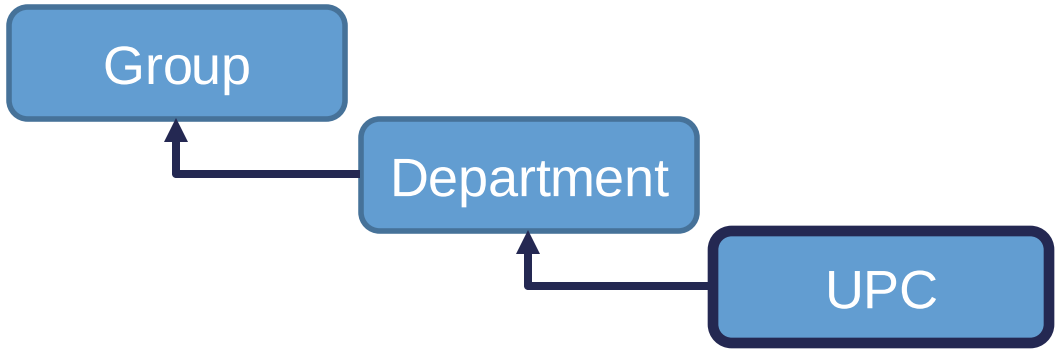
Reconciliation Approaches

Top-Down



Middle-Out

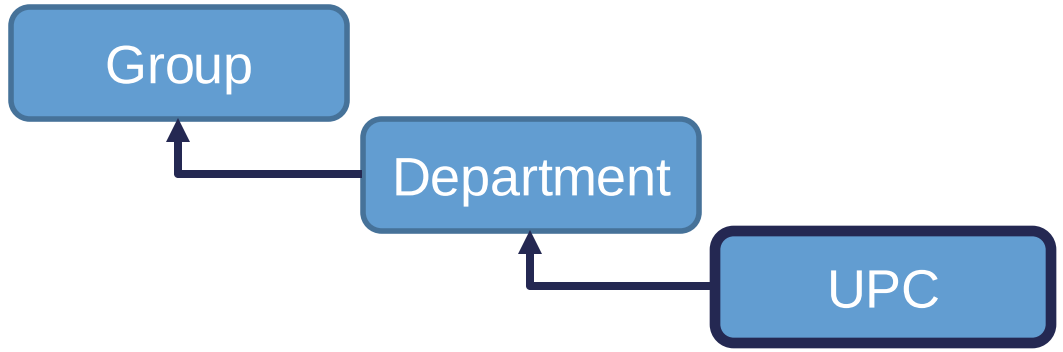
Bottom-Up



Bottom-up Approach

- Most common approach to hierarchical forecasting.
- Build a model for each series in the very bottom of the hierarchy structure.
- Add up the individual forecasts to build the hierarchy above it.

Bottom-Up



Bottom-up Approach

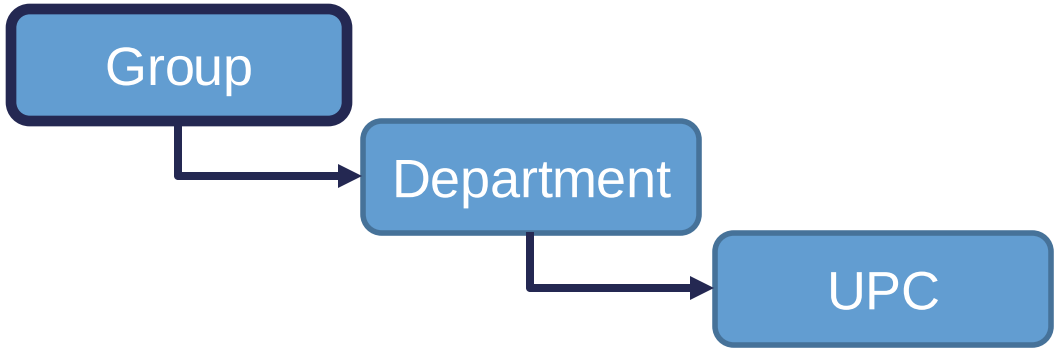
- Advantage:
  - NO LOSS OF INFORMATION!
- Disadvantage:
  - Very noisy data typically lies in the lowest hierarchy.
  - Potentially **MANY** models to build.

- All the detail is saved, but that means that you potentially have a lot of unneeded noise you are trying to model away.

Top-down Approach

- Simplest technique across all approaches.
- Build a forecast for the overall process and disaggregate this forecast down the tiers of the hierarchy.

Top-Down



Disaggregation Techniques

- There are a variety of different ways to disaggregate the forecasts in one tier down to the hierarchical tier below it.
- 3 Common Techniques:
  1. Average Historical Proportions
  2. Proportion of Historical Averages
  3. Forecasted Proportions

Average Historical Proportions

- Each proportion reflects the average of the historical proportions of the series a tier below relative to the total.

$$p_j = \frac{1}{T}\sum_{t=1}^{T}\frac{y_{j,t}}{y_t}$$

- Example:
  - On average, Department 1 historically makes up 34.8% of the total sales.

Example

- Example:
  - On average, Department 1 historically makes up 34.8% of the total sales.

	Year 1	Year 2	Year 3	Year 4
Department 1	$3,000	$5,000	$2,900	$4,000
Company	$10,000	$11,000	$9,500	$12,000
Proportion	**0.3**	**0.45**	**0.31**	**0.33**

Example

- Example:
  - On average, Department 1 historically makes up 34.8% of the total sales.

	Year 1	Year 2	Year 3	Year 4
Department 1	$3,000	$5,000	$2,900	$4,000
Company	$10,000	$11,000	$9,500	$12,000
Proportion	0.3	0.45	0.31	0.33

Average = 0.348

Proportion of Historical Averages

- Each proportion reflects the proportion of the historical averages of the series a tier below relative to the average total series.

$$p_j = \sum_{t=1}^{T} \frac{y_{j,t}}{T} \Big/ \sum_{t=1}^{T} \frac{y_t}{T}$$

- Example:
  - The proportion of Department 1's historical average makes up 35.1% of the average total sales.

Example

- Example:
  - The proportion of Department 1's historical average makes up 35.1% of the average total sales.

	Year 1	Year 2	Year 3	Year 4	Avg.
Department 1	$3,000	$5,000	$2,900	$4,000	**$3,725**
Company	$10,000	$11,000	$9,500	$12,000	**$10,625**

Example

- Example:
  - The proportion of Warehouse 1's historical average makes up 35.1% of the average total sales.

	Year 1	Year 2	Year 3	Year 4	Avg.
Department 1	$3,000	$5,000	$2,900	$4,000	$3,725
Company	$10,000	$11,000	$9,500	$12,000	$10,625

Proportion = 0.351

Disaggregation Techniques

- There are a variety of different ways to disaggregate the forecasts in one tier down to the hierarchical tier below it.
- 3 Common Techniques:
  1. Average Historical Proportions
  2. Proportion of Historical Averages
  3. Forecasted Proportions

Use time series models to forecast values of proportions over time!

Top-down Approach

- Advantage:
  - Simple to build as there are lower number of models needed.
- Disadvantage:
  - LOSS OF POTENTIAL VALUABLE INFORMATION!

- Potential loss of valuable information in the lower levels as our lower level forecasts are just proportions of upper level models that were easier to build.

Middle-out Approach

- Combines attributes from both bottom-up and top-down approaches.
- Aggregates up the tiers above with bottom-up approach.
- Disaggregates down the tiers with top-down approach.

Middle-Out

Middle-out Approach

- Tries to balance both approaches.
- Advantage:
  - Uses some of the detailed information, but not so detailed that it is too noisy.
- Disadvantage:
  - Doesn’t use ALL the information as disaggregate still takes place.