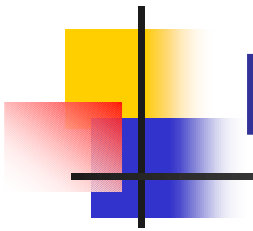
Best of Breed Modeling Techniques

Deb Campbell

January 12, 2001

Fingerhut Background

- Fifty year history
- Second in consumer catalog sales
- Offers general merchandise through catalogs, direct marketing and internet
- Wholly-owned subsidiary of Federated Department Stores
- 12,000 employees



Fingerhut Data

- 65 million customers on file, 7 million active customers
- Thousands of customer attributes:
  - Purchase history
  - Payment history
  - Promotion history
  - Demographics
- 80 unique catalogs mailed per year

Fingerhut Business Intelligence

- Staff of 30 supporting:
  - Fingerhut
  - Affiliates
  - Federated
- Analytical expertise
  - Statistics
  - Data Mining
  - DB Exploration
- Database Marketing Laboratory
  - Test software, methods
  - Encourage R&D

Modeling at Fingerhut

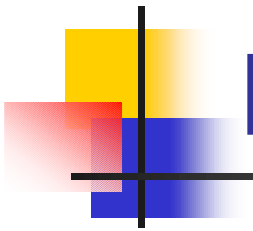
- Utilize past campaigns and results to predict outcome of future campaigns
- Customer level models estimate each customer's expected
  - Intent to purchase
  - Purchase amount
  - Intent to return
  - Profit
  - Other outcomes

What is BOB?

- A series of programs and software applications that replicate Business Intelligence's "best of breed" modeling processes

BOB Goals

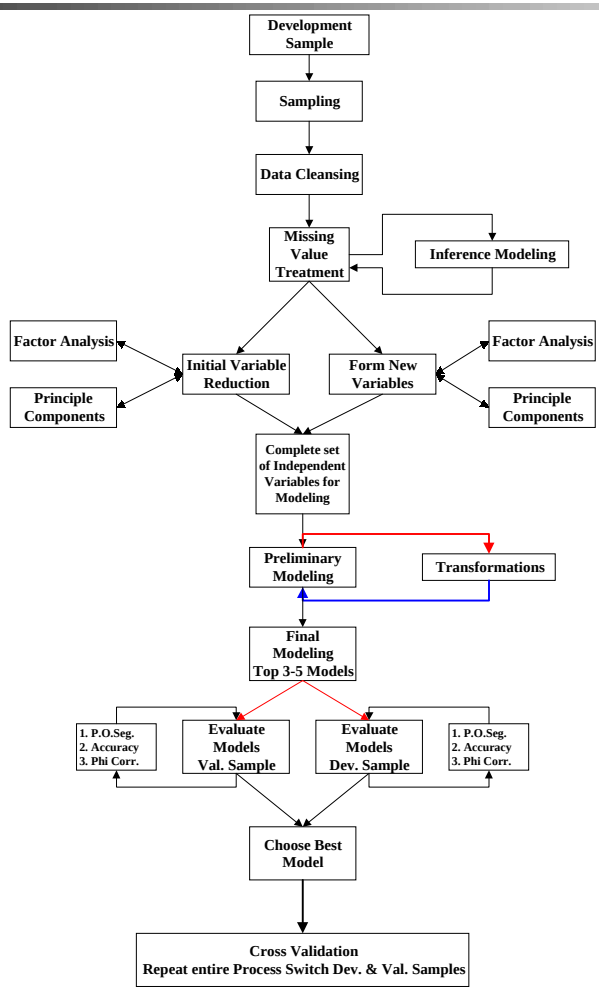
- Dynamically capture experience and expertise
- Build models efficiently
- Incorporate many details and decisions
- Provide standards
- Train new analysts, offer tutorial to all analysts



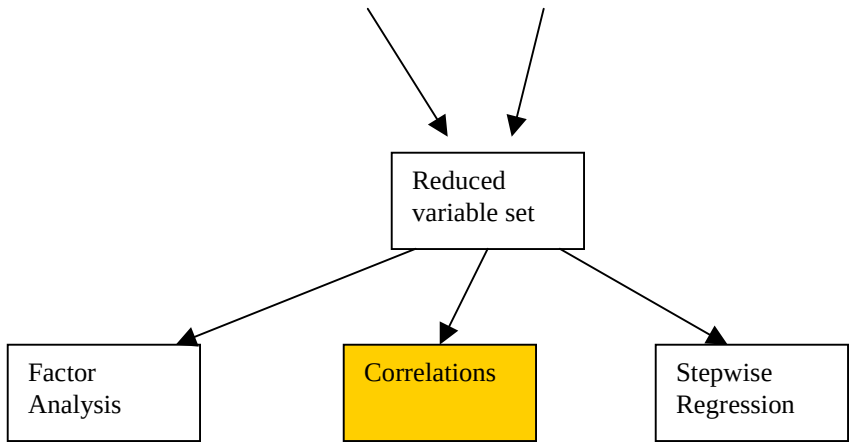
BOB Process

- Sampling
- Data Cleansing
- Missing Values
- Variable Reduction
- New Variables
- Modeling
- Model Evaluation

BOB Flow Chart



BOB Tutorial



BOB Tutorial (excerpt)

- ***Correlation***

Correlation is a statistical measure of the magnitude and direction of a relationship between two variables. The measure of correlation to use depends on the type of variables used (discrete, continuous, etc.). When both variables are continuous, the *Pearson correlation* is appropriate. Like all correlation measures, the Pearson estimates both the degree and the direction of linear relationship between any given two variables. Conceptually, the Pearson correlation can be depicted as follows:

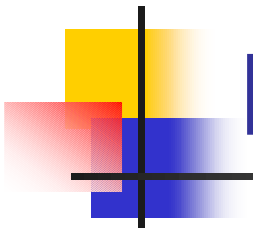
$$r = \frac{\text{degree to which } X \text{ and } Y \text{ vary together}}{\text{degree to which } X \text{ and } Y \text{ vary separately}}$$

In addition to correlation, which measures the magnitude and direction of relationships, the squared correlation (r^2) can also provide useful information. This squared value is known as the Coefficient of Determination and indicates the proportion of the change in *X* that can explain the change (variability) in *Y*.

When both variables are not continuous, other correlation measures are appropriate. When both variables are dichotomous (two categories) the ϕ (Phi) correlation is appropriate. If one variable is dichotomous and the other is continuous, then the *Point Biserial* correlation is used. If both variables are ordinal (ranked), then the *Spearman* correlation is used.

BOB Process: Variable Reduction

- Factor analysis
- Principle components
- Correlations
- Stepwise regression
- More



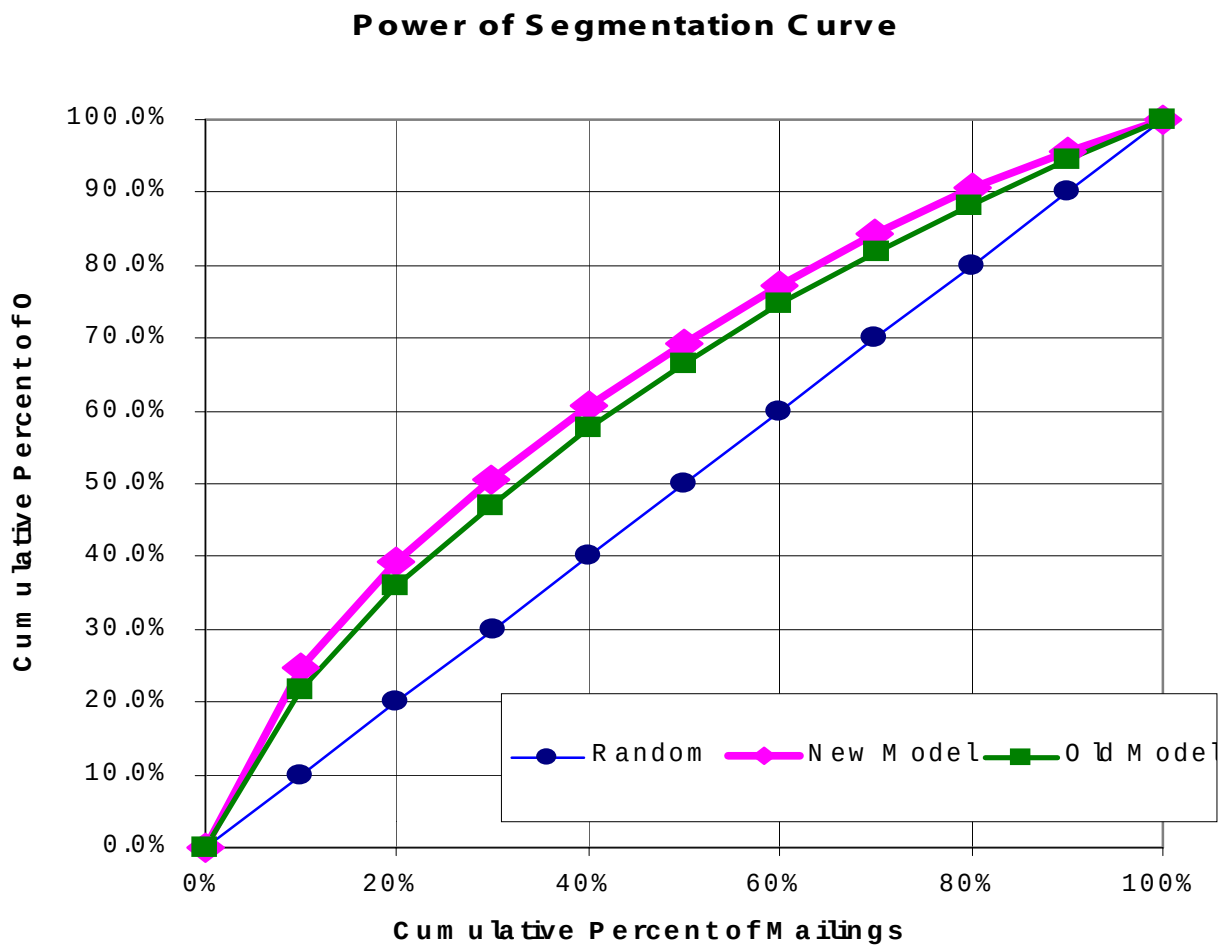
RAMP

- Tool built by IBM
- A candidate to incorporate into BOB
  - Use as variable reduction tool
  - RAMP’s feature selection
- Successful tests

ATM-SE (Advanced Targeted Marketing – Single Events)

- Tool built by IBM, in consultation with Fingerhut
- A candidate to incorporate into BOB
- Builds multiple models simultaneously
  - Finds appropriate sub-populations
  - Builds model for each sub-population
- Successful tests

BOB Process: Model Evaluation



BOB Efficiencies: Initial tests

Traditional Modeling

- 6-8 hours analyst time
- 1 week elapsed time
- 100% solution

BOB Modeling

- 10 minutes analyst time
- 20 minutes elapsed time
- 80% solution

BOB Applications

- Response model
- Returns model
- Sales model
- Credit Risk model
- Product preference
- Customer value
- Attrition/churn model
- Image map
- Clustering
- Product associations
- Fraud model
- Zipcode response

BOB Hardware/Software

- Windows NT
- Software independent
- Statistical languages (SAS, SPSS)
- Third party tools
  - Knowledge Seeker
  - Model 1
  - Quadstone
  - IBM tools

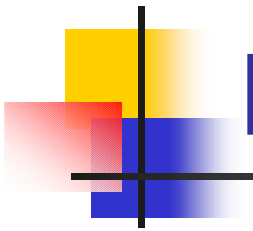
BOB Status

- BOB process patent pending as of 12/2000
- SPSS prototype built, active on SAS prototype
- Using BOB internal to Fingerhut
- Will issue BOB challenge to affiliates
- Continually updating content, functions, usability

BOB Potential Partnerships

- Fingerhut has modeling expertise, but not product development nor sales expertise
- Consulting engagements
- Product development
- Internet-based modeling tool

BOB Benefits



- Increase revenue for Fingerhut, affiliates, Federated
- Improve BI analyst productivity
- Save on computing costs
- Potentially increase outside revenue
  - Joint consulting projects
  - Joint product development

Best of Breed Modeling Techniques

Questions?