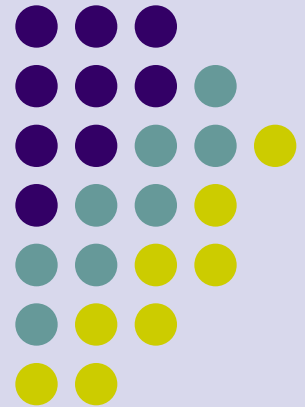


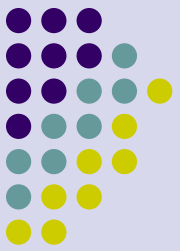
Principles of Data Mining: Numbers *are* Your Friends

Kevin Whorton,
Whorton Marketing & Research



I don't even see the code. All I see is blonde, brunette, redhead ...
- The Matrix

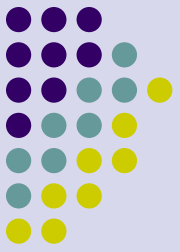
- Data Mining as a Discipline



In existence many years

- **Fed heavily by dot-com/e-commerce/online**
 - Suddenly we have tons of data pre-captured
- **Sometimes a derogatory term**
 - Social science: “hidden patterns”=no up-front hypothesis
- **Critical to business success**
- **In association marketing & membership:**
 - Document behavior (transactions)
 - Easy linkages to attitudes (surveys)
 - Evidence-based decision-making
 - Basis for experimentation

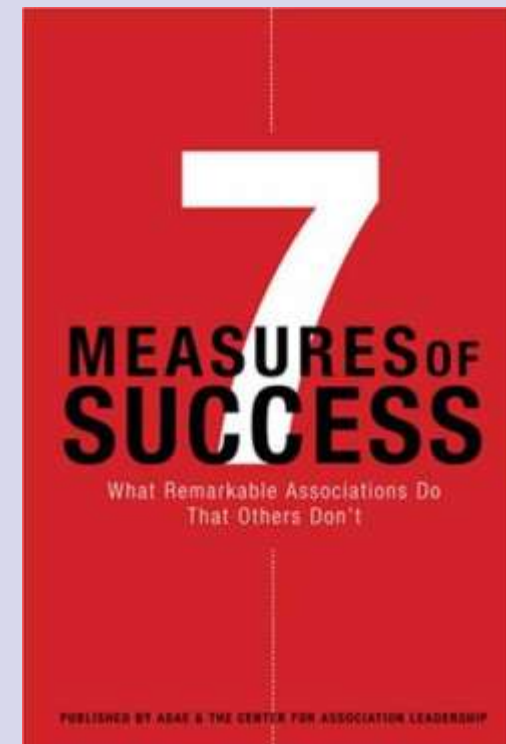




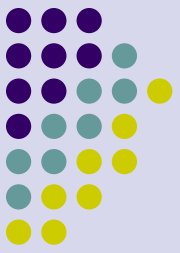
- Why Data Mining Is Critical

In many associations our programs are:

- **Often ruled by perceptions**
 - Anecdotal evidence of success=ok
 - Quality judged by design/copy, not audience, usage, results
- **Metrics are often flawed/insufficient**
 - Poor measures of share of market, mind or wallet
 - Sales tracking rates low
 - Weak/no early warning systems
- **Data mining addresses all of these**
 - ASAE recognized “Data Driven Decision Making” as one of the 7 Measures



- Data Hygiene

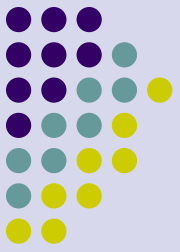


GIGO principle: Garbage in/Garbage Out

- **Maintaining or making data clean is critical**
 - Capturing right the first time is best:
Or, periodic “sweeps” to clean data
 - Spreadsheets, spell checkers, global search & replace, and re-importation
 - Correct & standardized spelling, complete more data fields
 - Facilitates analysis, boosts deliverability and response rates



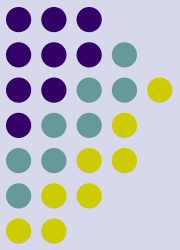
- Make the Most of Your Tools



Case studies:

- **PPAI**: Data cubes quantify value of membership each renewal cycle
 - Numeric benefits, FedEx, other affinity programs
- **TMA**: Years of effective use of in-house network's MS SQL
 - Drives county level education & membership development
- **NDIA**: Using BI/AMS/e-marketing
 - Doubled revenue/membership





- Examples: Nonprofit Donor Fundraising

Information-rich environments very common

- **Data sharing through Target Analysis Group**

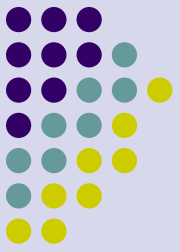
- Quarterly industry-wide & sector metrics
- \$20,000 annual benchmarking service: 2-day meetings with competitors, reports showing performance of each by name

- **DMA-NF presentations**

- Sharing packages, techniques, results
- Constant internal testing
- Library of results



Packages:	A	B	C
# Mailed	364,945	150,000	50,000
% Response	0.71%	1.23%	0.77%
Average Gift	\$36.48	\$23.73	\$55.52
Rev/M	\$257.26	\$292.99	\$424.76
CPDR	\$1.38	\$1.54	\$0.93



- Managing to/Creating Your Metrics

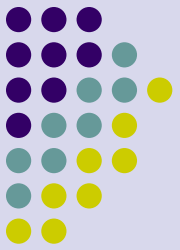
Recent listserv discussion/illustration: how do you project/track success?

● Meetings oriented example

- Aggregate plans, traffic report, actual response data, and evaluations
- Provides audience size, known (tracked) response, inferred response
- Paint the entire picture: how many registrations do you expect; how many do you attain?
- What is your ROI overall and by channel?
- How would/should this drive your budget size, allocation decisions?

Effort	Audience	Resp rate	Regist.	Gr. Rev.	Cost
Mail I: Week 18	2,000	1.00%	20	\$11,900	\$3,000
Email I: Week 16	7,980	0.50%	40	\$23,800	\$100
Mail II: Week 13	1,940	2.50%	49	\$29,155	\$3,000
Email II: Week 10	8,000	0.50%	40	\$23,800	\$100
Email III: Week 7	8,000	0.50%	40	\$29,000	\$100
Overall:	27,920	NA	189	\$117,655	\$6,300

- Applying Structured Research



Most surveys are disconnected from central data

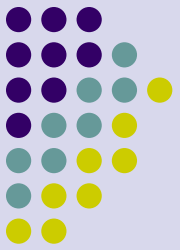
- Even when linked to AMS (i.e. Informz): little/no integration of survey results
- Unfortunate: surveys can re-populate demographics, business characteristics (overlay data)

- **Market research validates & explains marketing results**

- Why & how vs. what
- Understand perceptions/motivations behind action
- Understand behavior/attitudes by segment
- Test/assess feasibility of concepts prior to launch
- Qualitative research to define positioning, drive messaging
- Competitive intelligence/environmental scans



- AMS As Frenemy



Key roles of Assn Management Systems

- Good tool for collecting & storing data
- Some report capability: often strains to be all things to all people
- Other better tools allow you to see forest for the trees

The screenshot shows a web-based interface for managing an event. The main content area is titled 'AVECTRA 09 EVENT' and includes the following sections:

- Registration Information:** Shows 'begins on: 10/30/2009' and 'ends on: 11/02/2009'. There are input fields for 'pre reg:', 'early:', 'standard:', and 'late:'.
- Registration Options:** Includes checkboxes for 'registration required?' (checked), 'wait-list allowed?' (unchecked), 'abstract?' (unchecked), and 'automatic registration from wait-list?' (unchecked). A note '(Space Permitting)' is next to the last option.
- Miscellaneous Information:** Shows 'code/type: AV09 / Conference' and 'category: NCCAE'. A red arrow points to an 'EDIT EVENT INFO' button with the text 'Click on Edit Event Info'.
- Publishing & Communication:** Includes 'post to web:', 'time zone: (GMT-05:00) Eastern Time (US & Canada)', and 'confirmation:'.
- Navigation Icons:** A row of icons for 'individual registration wizard', 'group registration wizard', 'quick group registration wizard', 'sponsor', 'search participants', 'list registrants', and 'find registrants'.

On the right side, there is a 'STATISTICS' panel with the following data:

- goals:** registration: revenue: 0.00
- attendance:** projected: guaranteed: capacity:
- counts:** registered: 79, on wait list: 0, available: remaining:

At the bottom right, there is an 'EVENT STATISTICS' section with a list: 'Total Reg Types', 'Total Board Members', and 'Total Picked Up'.



- SAS: Basic BI/Data-Mining Tool

Easy to use tool:

- Import data
- Run cross-tabulations, correlations
- Data-step programs to merge data
- Easy queries:
- Customer, member analysis
- \$1600/year software license

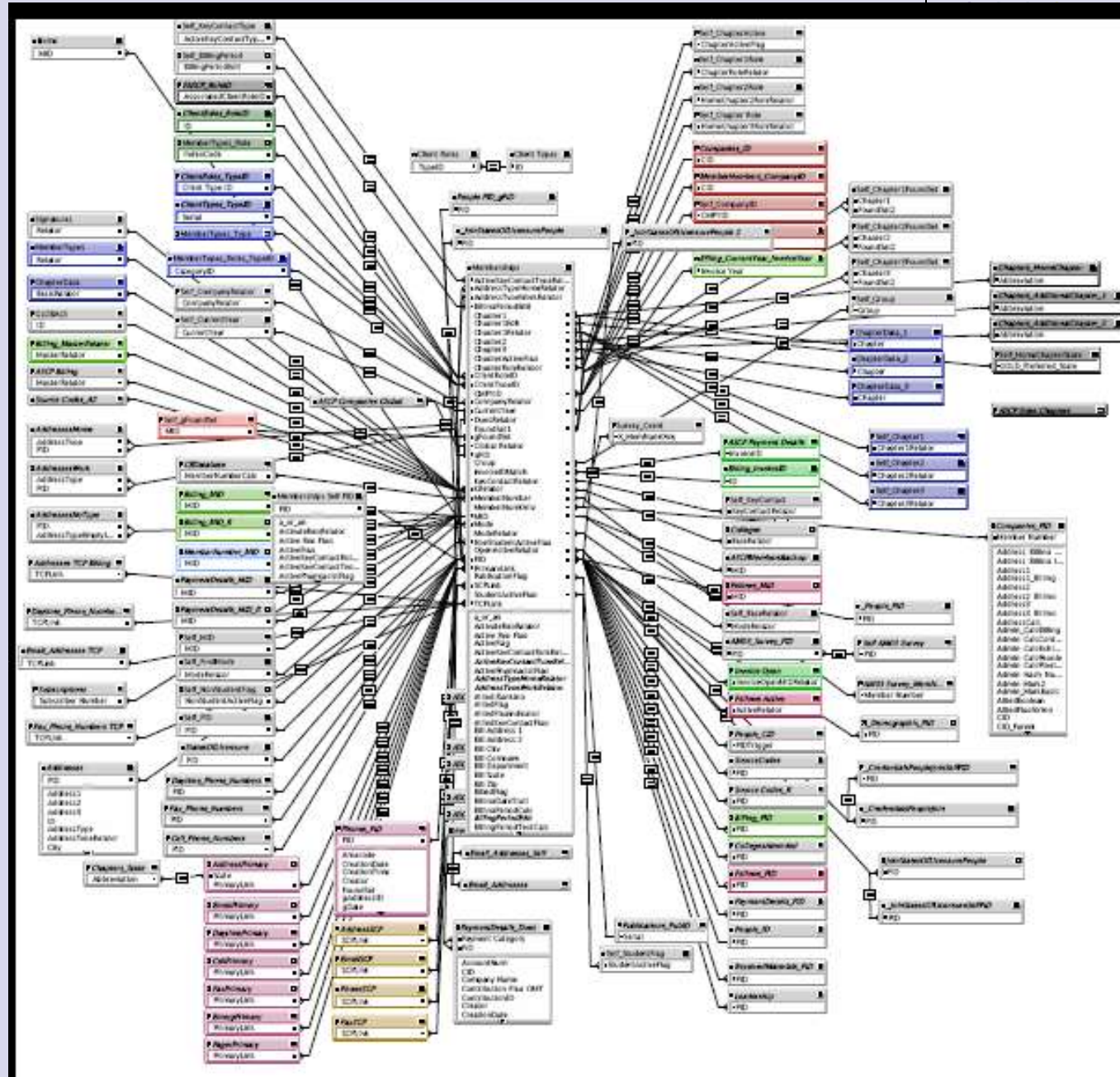
The screenshot displays the SAS 9.1.3 Service Pack 4 interface. On the left, a file explorer window shows a directory structure with files like 'Aaas', 'Alongit', 'Mem', 'Mem05b', 'Mem05d', 'Mem06b', 'Mem06d', 'Mem07b', 'Mem07d', 'Mem08b', 'Mem08d', 'Mem09b', 'Mem09d', 'Mem5', and 'Survey'. The main window is split into two panes. The top pane, titled 'Log - (Untitled)', shows system messages including copyright information for SAS Institute Inc. (2002-2003), SAS 9.1 (TS1M3) license details for WHORTON MARKETING & RESEARCH, and initialization statistics: real time 3.26 seconds and CPU time 0.67 seconds. The bottom pane, titled 'means2.sas *', contains SAS code:

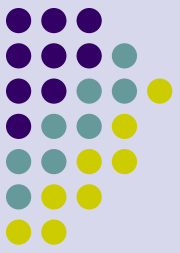
```
libname x 'c:\proj\clients\aaas';  
  
proc sort data=i.ifma; by alsfr;run;  
proc means data=i.ifma noprint;  
  var sfratio; output out=one mean(sfratio)=a;by alsfr; run;  
proc print data=one;run;
```



- Tech Staff As Frenemy

- Actual database structure
- Provided in response to request for data dictionary and glossary
- Many systems are poorly documented
- IT's role is to maintain at a reasonable cost
- Mission-critical, just not best environment for data mining



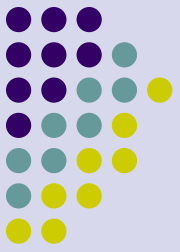


- My Basic Approach

1. Define goals: specific applications and lessons to learn
2. Assess data resources: current/past databases, research report, new primary collection
3. Download AMS data, reconfigure, transform, clean, and load into SAS
4. Begin running queries: share preliminary results
5. Expand the discovery process into spreadsheets, narrative reports
6. Refine goals/identify new hypotheses
7. Complement existing data with newly-collected, share with AMS
8. Apply to marketing tests, segmentation, personalization
9. Continue process in traditional cycle



- Illustration: Key Areas of Inquiry



1. Member life cycle: join, engagement, attrition
2. Basic item/campaign profitability
3. Market testing/feasibility assessment
4. Campaign success measures/assessing channel impact
5. Cross-selling/conversion rates/indicators
6. Measuring member and customer value
7. Event analysis: frequency/stability of purchase/actions
8. Response modeling: linear regression/ANOVA
9. Assessing price sensitivity/willingness to pay
10. Basic segment analysis: descriptive cross-tabulations
11. Audience cluster analysis: psychographics/demographics
12. Exception reporting/early warning for defection