



**THE
POWER
TO KNOW[®]**

SAS[®] Enterprise Miner[™]

Sue Walsh
Higher Education Consultant
Education Division

Scenario

- A supermarket is beginning to offer a line of organic products. The supermarket's management would like to determine which types of customers are likely to purchase these products.
- The supermarket chooses to test market this new line of products in several of its stores.
- The supermarket has a customer loyalty program. As an initial buyer incentive plan, the supermarket provided coupons for the organic products to all of their loyalty program participants and have now collected data that includes whether these customers have purchased any of the organic products.

The Data

Name	Description
CUSTID	customer loyalty identification number
GENDER	M = male, F= female, U = unknown
DOB	date of birth
EDATE	date the data was extracted from the daily sales data base
AGE	age, in years
AGEGRP1	grouped values of age
AGEGRP2	alternate grouped values of age
TV_REG	television region
NGROUP	neighborhood group

continued...

The Data

Name	Description
NEIGHBORHOOD	type of residential neighborhood
LCDATE	loyalty card application date
LTIME	time as a loyalty card member
ORGANICS	number of organic products purchased
BILL	total amount spent
REGION	geographic region
CLASS	customer loyalty status: tin, silver, gold, or platinum
ORGYN	Organics purchased? 1 = Yes, 0 = No
AFFL	affluence grade on a scale from 1 to 30

Data Issues

- Dates converted to length of time
- Duplicate fields – neighborhood, region, and age
- Two target variables



Demonstration

Conclusion

- The characteristics of those most likely to purchase organic products are:
 - Higher affluence
 - Female
 - Middle age
- These products might do well at stores where many of the customers have these characteristics.