

**THE
POWER
TO KNOW.**

Predictive Modeling in Retail

Jim Godfrey – SAS Institute (Canada)

Objectives

- The overall objectives of the Retail Predictive Modeling activities:
 - Build three predictive models to support the development and execution of three marketing campaigns.
 - Build the process required to extract the required data for model development and for ongoing scoring.
 - Knowledge transfer with the Modeling resource to make sure they are able to repeat the process

Predictive Models

- **Campaign Response Model** – this model predicts the likelihood that a customer responds to a specific campaign by purchasing a products solicited in the campaign. The model also predicts the amount of the purchase given response.
- **Cross Shop Model** – This model predicts the likelihood that a customer who has not made any purchases from a specific department within the company in the previous 18 months makes a purchase from this department within the target month. In addition, the model predicts the amount of the purchase.
- **Four-month Reactivation Model** – This model predict the likelihood that a customer who has not made any purchases from the company at all in the previous four months spontaneously makes a purchase within the target month. Similar to the other two models, this model also predicts the amount of purchase given reactivation.

Data Sources and Model Inputs

- The model inputs include variables created from the following sources within the SAS data mart:
 - Customer Transactions
 - Transaction Tender
 - Distinct Counts
 - Customer Promotion History
 - Customer Characteristics
- Counts and Amounts by Division/Class
- Counts and Amounts by transaction tender
- Past promotions
- Store visits
- Etc.

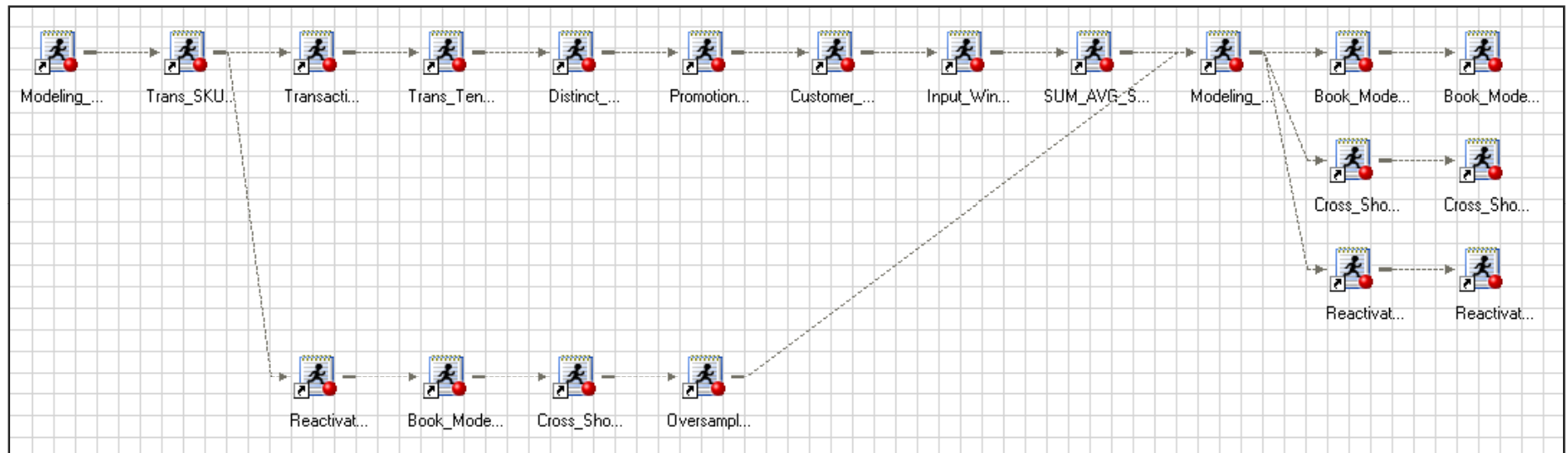
Sliding Windows Concept

Feb 05	Mar 05	Apr 05	May 05	Jun 05	Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07	Sep 07	Oct 07	Nov 07	
																							Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07	Sep 07	Oct 07	Nov 07
																							Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07	Sep 07	Oct 07
																							Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07	Sep 07
																							Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07	Aug 07
																							Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07	Jul 07
																							Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07	Jun 07
																							Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07	May 07
																							May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07	Apr 07
																							Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07	Mar 07
																							Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07	Feb 07
																							Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06	Jan 07
																							Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06	Dec 06
																							Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06	Nov 06
																							Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06	Oct 06
																							Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06	Sep 06
																							Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06	Aug 06
																							Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06	Jul 06
																							Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06	Jun 06
																							Jun 05	Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06	May 06
																							May 05	Jun 05	Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06	Apr 06
																							Apr 05	May 05	Jun 05	Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06	Mar 06
																							Mar 05	Apr 05	May 05	Jun 05	Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06	Feb 06
																							Feb 05	Mar 05	Apr 05	May 05	Jun 05	Jul 05	Aug 05	Sep 05	Oct 05	Nov 05	Dec 05	Jan 06

Additional Summary Model Inputs

- Overall Average
- Average for month 1-3 of the window
- Average for month 4-6 of the window
- Average for month 7-9 of the window
- Average for month 10-12 of the window
- Standard Deviation for month 1-3 of the window
- Standard Deviation for month 4-6 of the window
- Standard Deviation for month 7-9 of the window
- Standard Deviation for month 10-12 of the window
- Trend variable over all 12 months in the window (SLOPE of a regression of the variable against month)
- Proportion of visits by location (store)
- Proportions of fees, markdowns, sales, returns and discounts

Data Extraction Process

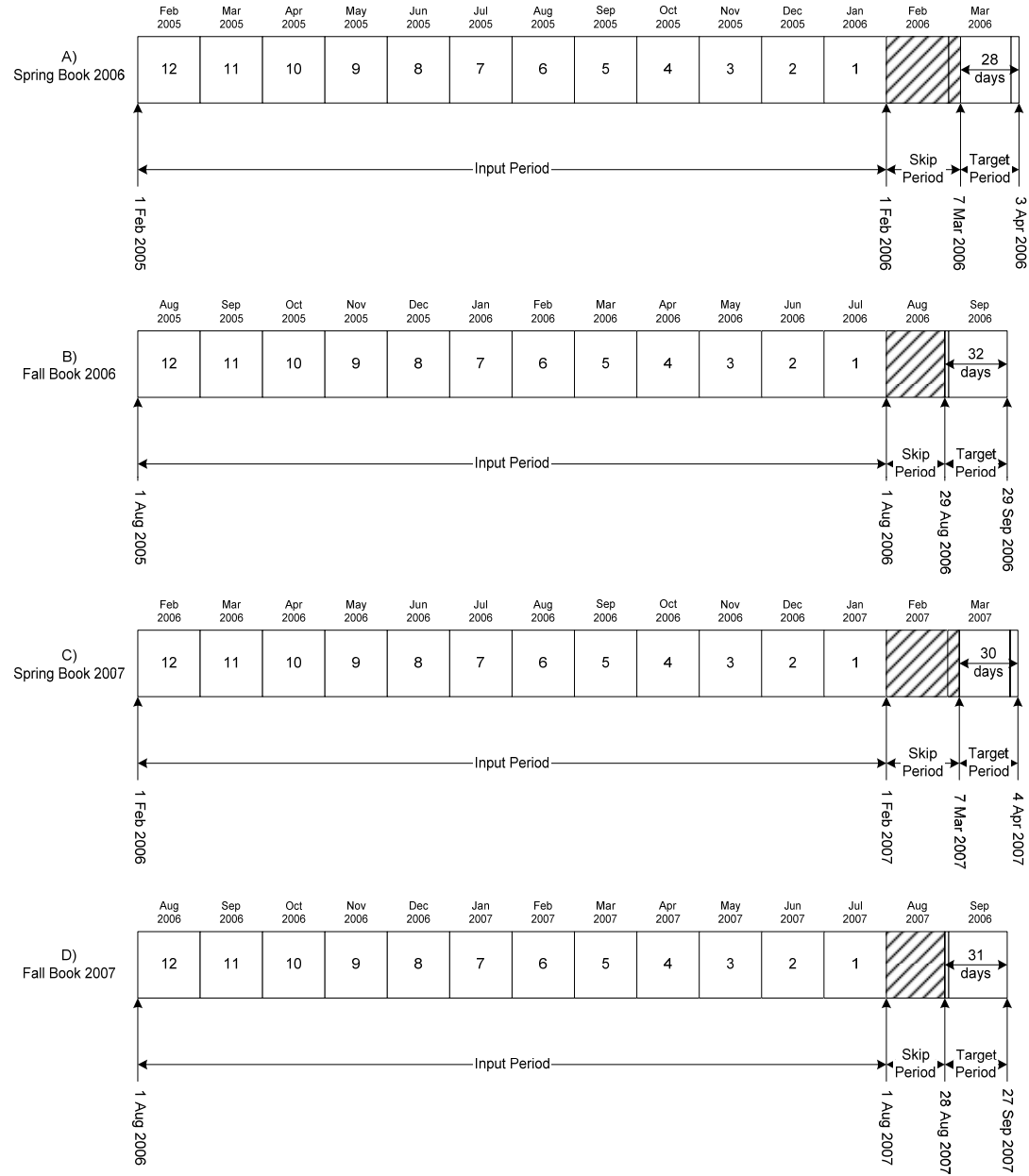


Campaign Response Model

- The purpose of this model is to predict those customers that are likely to respond to a specific campaign; irrespective if it is a Fall or Spring campaign. This model will be used to score and select a population to receive the Fall campaign from the company.
- Response to the campaign is implied through observing customer's shopping behavior during the response period using the following conditions:
 - The customer has shopped and purchased items identified as solicited items within the allocated response period OR
 - The customer has shopped and purchased *any* items within the allocated response period

Campaign Response Model

Timeframes



Campaign Response Model

Timeframes

A)	12	11	10	9	8	7	6	5	4	3	2	1		← 28 days
B)	12	11	10	9	8	7	6	5	4	3	2	1		← 32 days
C)	12	11	10	9	8	7	6	5	4	3	2	1		← 30 days
D)	12	11	10	9	8	7	6	5	4	3	2	1		← 31 days



A)	12	11	10	9	8	7	6	5	4	3	2	1		Target
B)	12	11	10	9	8	7	6	5	4	3	2	1		Target
C)	12	11	10	9	8	7	6	5	4	3	2	1		Target
D)	12	11	10	9	8	7	6	5	4	3	2	1		Target

Campaign Response Model – Population and Targets

- Campaign Model Population
- Campaign Model Target Definitions:
 - **English Definition:** A prospect, after receiving the solicitation, shops at the store and makes a purchase within the observational period.
 - **Data Definition:** if TRANS.BUSINESS_DATE between (*start_date*, *end_date*) AND TRANS.SALES_TYPE = 'Sale' and TRANS.SALES_AMOUNT > 0 then t = 1, else t = 0.
 - Target = MAX(t) by CUSTOMOER_ID
 - **English Definition:** A prospect, after receiving the campaign solicitation, shops at the store and makes a purchase of an item featured in the campaign within the observational period.
 - **Data Definition:** if TRANS.BUSINESS_DATE between (*start_date*, *end_date*) AND TRANS.SALES_TYPE = 'Sale' and TRANS.SALES_AMOUNT > 0 AND SKU_EXTENDED_DETAILS.BOOK_CODE = '*specific campaign code*' then t = 1, else t = 0.
 - Target = MAX(t) by CUSTOMOER_ID

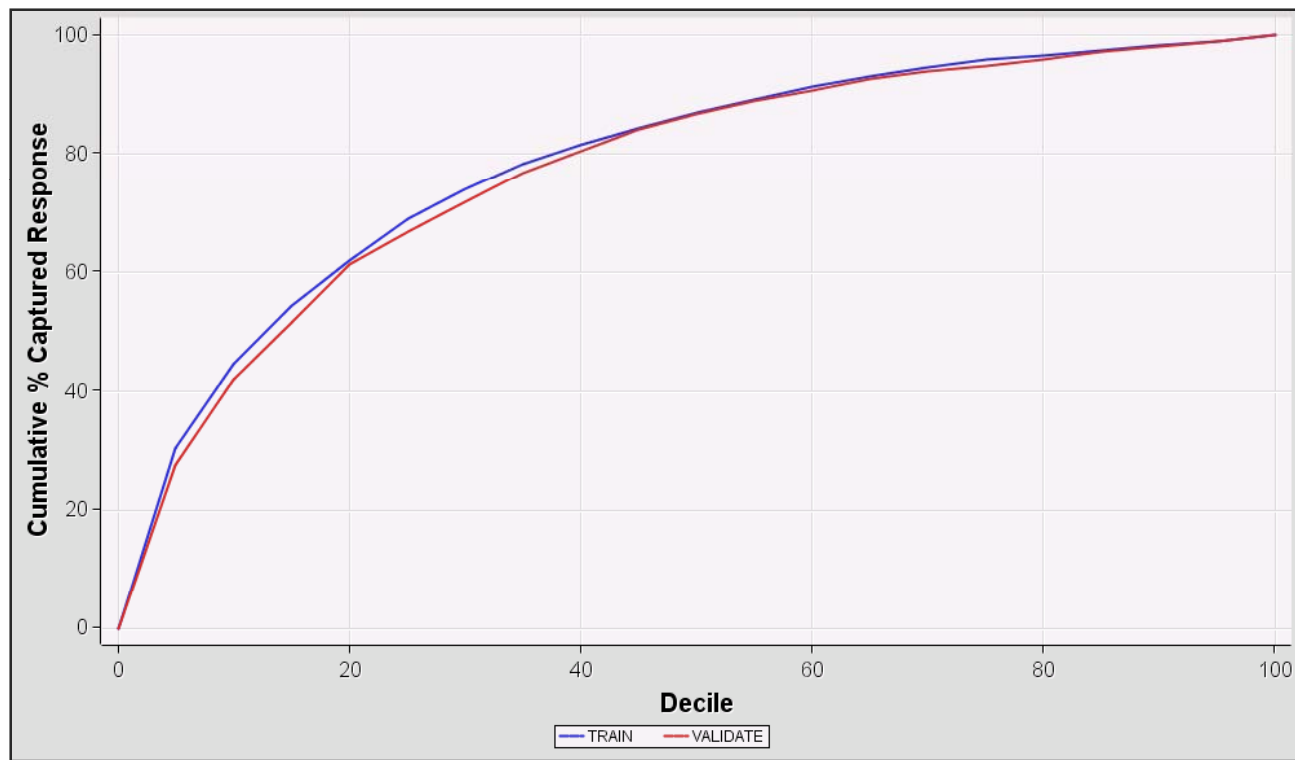
Campaign Response Model – Preliminary Variable Selection

- The initial set of potential model inputs was refined through a semi-automatic process:

Selection Step	Variables Kept	Variables Dropped
Start	3,890	-
Unary	3,474	416
Sparse	2,179	1,295
Redundant	503	1,676

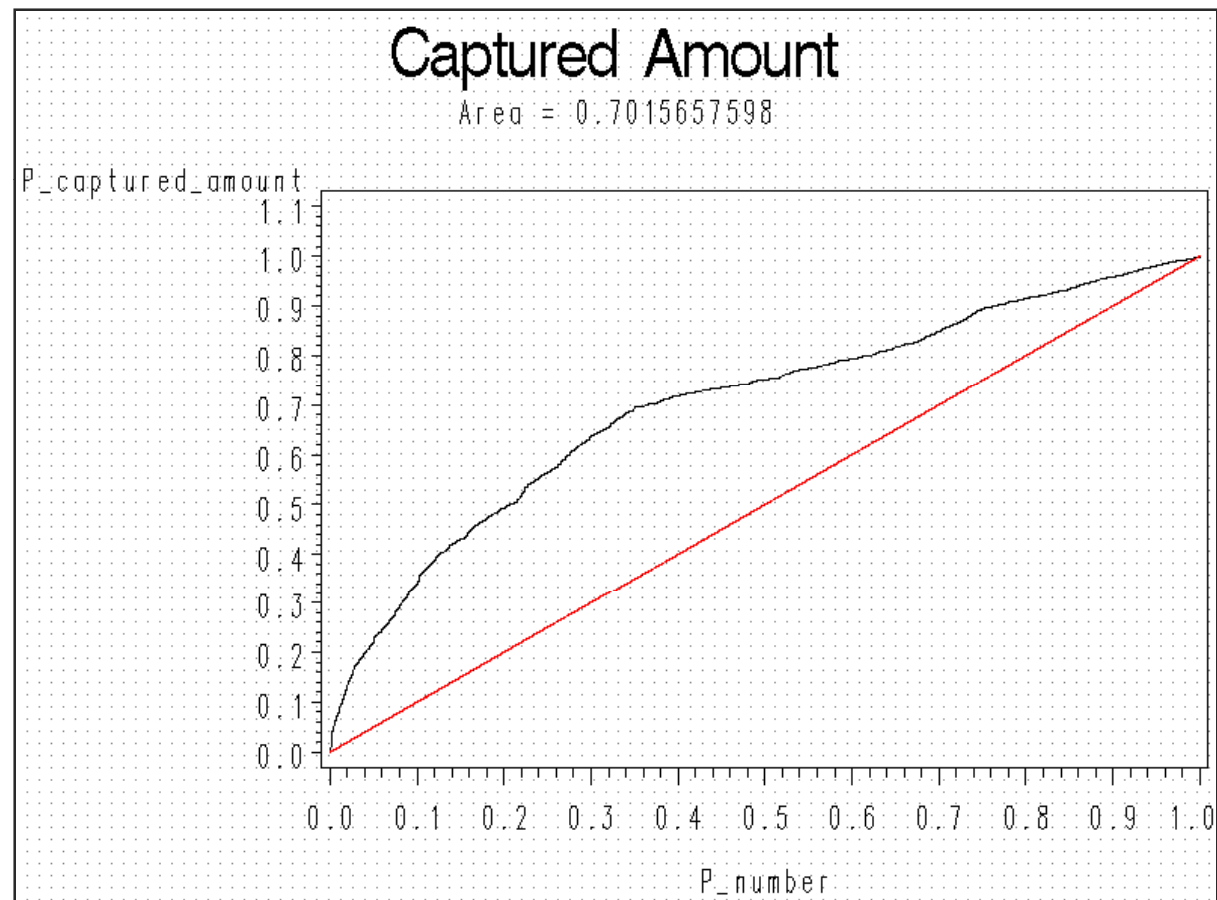
Campaign Response Model – Results

- Binary Target - % Captured Response



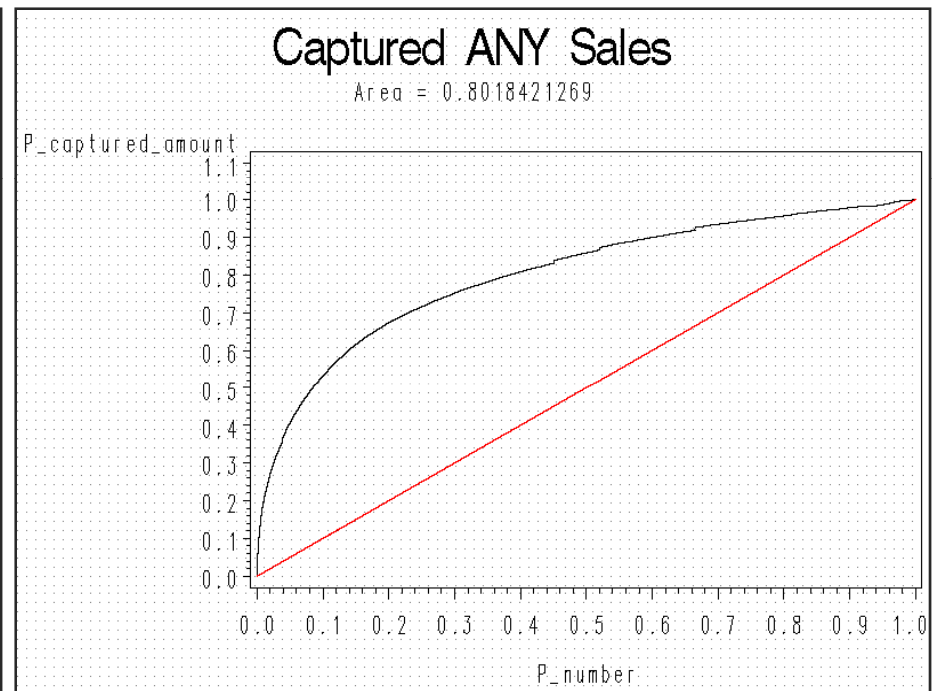
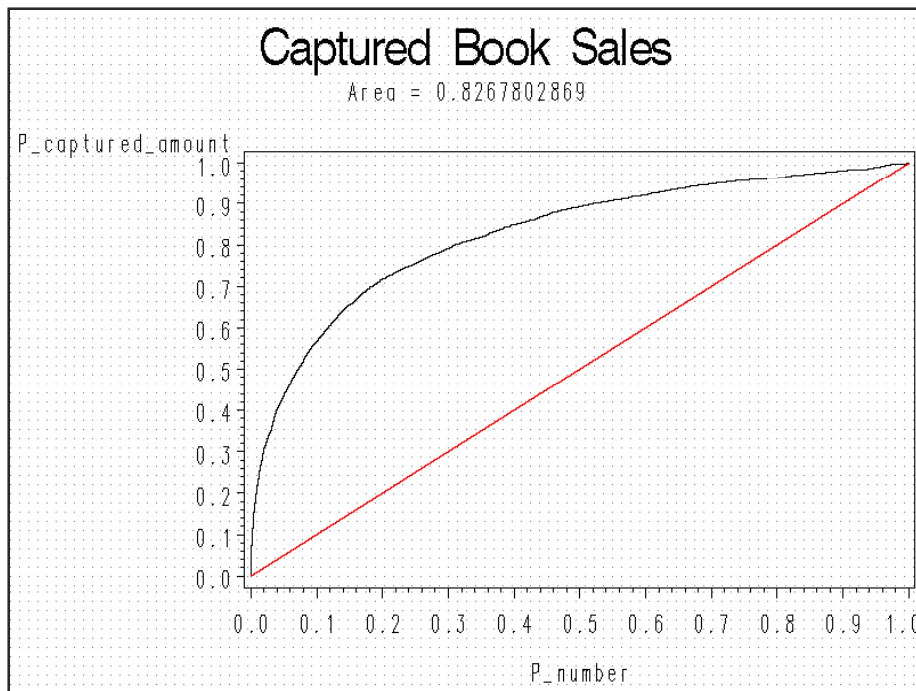
Campaign Response Model – Results

- Amount Target - % Captured Response



Campaign Response Model – Results

- Campaign Sales vs. Any Sales



Campaign Response Model – Results

Binary Model

Effect	DF	Chi-Square	Pr > ChiSq
AVG_BOOK_NUM_7to9	1	37.9354	<.0001
AVG_NUM_Promo_10to12	1	26.1618	<.0001
AVG_NUM_Promo_GWP_1to3	1	25.1282	<.0001
AVG_VISIT	1	19.1816	<.0001
<u>IMP_HR_Amex_Card_Type</u>	2	25.9353	<.0001
LOG_STD_DEP_056_NUM_1to3	1	15.2266	<.0001
LOG_STD_DIV_70_NUM	1	11.7157	0.0006
PCTL_AVG_BASIC_NUM	2	20.3109	<.0001
<u>PCTL_AVG_TOT_Sales_AMT</u>	3	34.2933	<.0001
PCTL_AVG_VISIT	3	24.9743	<.0001
PCTL_LOG_AVG_DIV_50_NUM_10to12	1	27.0523	<.0001
PCTL_SLOPE_BASIC_NUM	3	26.5082	<.0001
PCTL_SLOPE_DEP_019_NUM	3	32.3317	<.0001
PCTL_SLOPE_LOC_71_AMT	1	20.453	<.0001
PCTL_SLOPE_PT_14_NUM	1	25.0713	<.0001
PCTL_STD_Distinct_SKU_NUM_4to6	2	20.5802	<.0001
<u>PCTL_STD_Line_NUM</u>	3	16.7675	0.0008

Amount Model

NAME	LABEL	NRULES	IMPORTANCE
<u>AVG_TOT_Sales_AMT</u>	AVG of Total Sales Amount	6	1
PCTL_SLOPE_DEP_019_NUM	Transformed: Trend Slope for Line Num for (50) COSMETICS->(019) COSMETICS	1	0.46479
AVG_DEP_019_NUM_7to9	AVG of Line Num for (50) COSMETICS->(019) COSMETICS month 7 to 9	2	0.34613
AVG_DIV_50_NUM_10to12	AVG of Line Num for (50) COSMETICS month 10 to 12	1	0.18103
<u>SLOPE_NUM_Promo_GWP</u>	Trend Slope for Number of Gift with Purchase Promotions	1	0.14395
LOG_P_PT_14_AMT_1	Transformed: : 1	1	0.13112
PWR_P_Markdown_2	Transformed P_Markdown_2	1	0.11863
SLOPE_DIV_42_NUM	Trend Slope for Line Num for (42) ACCESSORIES	1	0.11198
AVG_VISIT	AVG of Visits to Holt Renfrew	1	0.10831
LOG_STD_TOT_Markdown_AMT_4to6	Transformed: STD of Total Markdown Amount month 4 to 6	1	0.07276

Questions



**THE
POWER
TO KNOW®**