



SPSS

## > Optimizing Case Management with Predictive Tax Compliance

<b>SPSS</b>	<b>CGI</b>
<b>Benjamin Chard</b>	<b>Ted London</b>
Senior Solution Engineer bchard@spss.com	Executive Consultant Tax and Revenue Management <a href="mailto:ted.london@cgi.com">ted.london@cgi.com</a>

SPSS

## > Predictive Tax Compliance

	Canada Revenue Agency	Agence du revenu du Canada
		
		

© 2007 SPSS Inc. 2



### > Where Does Predictive Tax Compliance Fit

Existing Data

Operational Setting

© 2007 SPSS Inc. 3

### > Data Warehouse Solution

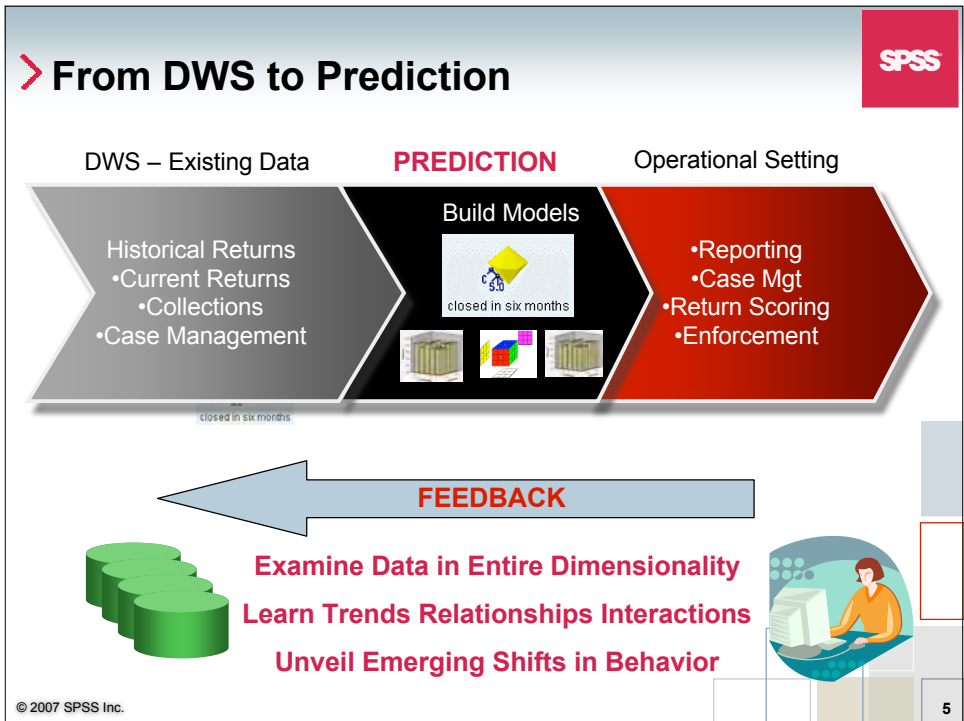
DWS – Existing Data

- Historical Returns
- Current Returns
- Collections
- Case Management

Operational Setting

- Reporting
- Case Mgt
- Return Scoring
- Enforcement

© 2007 SPSS Inc. 4

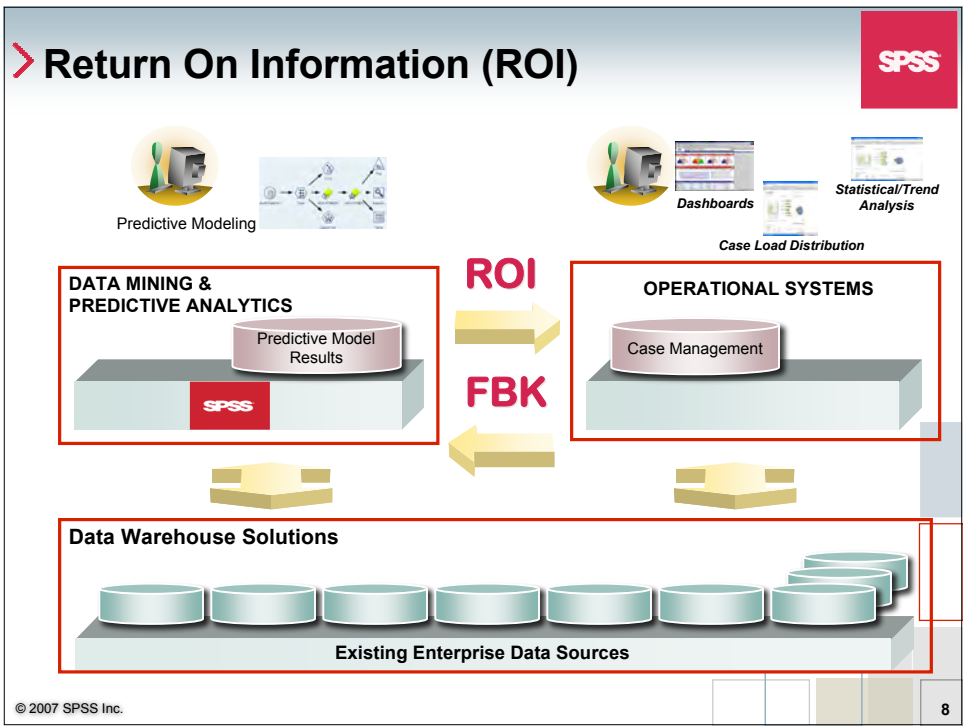


- ## > Predictive Modeling Scenarios
- Revenue: better return on limited resources
    - Audit Outcome
    - Predict \$per audit hour
    - Collections – Treatment Streams, Collection Activities
  - No-Change: Decrease the number of false-positive investigations
    - Hit Ratio
    - Self Cure
  - Proactive Approach / Reaction Time
    - Coverage
    - Legislative Concerns
    - Taxpayer Attrition
    - Quicker Enforcement
    - Emerging Trends
- © 2007 SPSS Inc. 6



**SPSS**

➤ **We are talking about Return on Information (ROI)**





**SPSS**

➤ **SPSS: Building Predictive Models**

**DATA MINING & PREDICTIVE ANALYTICS**

Predictive Model Results

SPSS

C:\Documents and Settings\mfalkowski\Desktop\New Presentations\Raytheon\EFM\theSafe\index.htm - Microsoft Internet Explorer

File Edit View Favorites Tools Help Address C:\Documents and Settings\mfalkowski\Desktop\New Presentations\Raytheon\EFM\theSafe\index.htm Go

Back Forward Stop Search Favorites Media Google Search Check

**SPSS**

➤ **SPSS Inc. is the world's leading provider of predictive analytics software and solutions.**

[About SPSS](#)  
[Current Customers](#)  
[What is Predictive Analytics?](#)  
[Predictive Analytics Architecture](#)  
[The Demo](#)

**About SPSS**

- Founded in 1968
- NASDAQ-listed (SPSS)
- 37+ year heritage as an innovator in predictive analytic technologies.
- More than 1,200 employees
- Operations in more than 60 countries
- Over \$220 million in revenues
- More than 250,000 customers

**SPSS Industries**

- US Military
- US Intelligence
- Homeland Security
- Banking & Financial Services
- Higher Education
- Healthcare & Insurance
- Manufacturing
- Market Research
- Retail / Consumer packaged goods
- Telecommunications

**SPSS Facts**

- Used in every branch of US Military and US Intelligence Community
- More than 95% of FORTUNE 1000 and more than 95% of FORTUNE 100 are SPSS customers
- More than 90% of top universities use software from SPSS
- All 50 U.S. state governments and every U.S. cabinet-level department uses SPSS analytics
- Formal partnerships with majority of System Integrator Community
- Twenty-four of the 25 top global market research organizations rely on SPSS
- More than half of the telecommunications companies on the Forbes 500 and S&P 500 are SPSS customers
- The top 10 global banks use software from SPSS

start My Computer 2:51 PM



## SPSS

### > Predictive Analytics

- Three classes of data mining algorithms:
  - Associate "Patterns"**

What events occur together? Given a series of actions; what action is likely to occur next?
  - Cluster "Differences"**

Group cases that exhibit similar characteristics.
  - Predict "Relationships"**

Predict who is likely to exhibit specific behavior in the future.

Data Mining

© 2007 SPSS Inc. 11

## SPSS

### > Concept Extraction

"Mr. Smith aka Mr. Ahmed was seen on the corner of Church St. and Magnolia Ave. on Nov 13<sup>th</sup>"

**Bag of « Words » extraction**

Mr. Smith  
aka  
Ahmed  
was seen  
on the  
corner  
of  
Church  
St.  
Etc.

**Expressions extraction**

Mr. Smith  
was seen  
Mr. Ahmed  
corner  
Church St.  
Magnolia Ave.  
Nov 13<sup>th</sup>

**Named Entities extraction**

Mr. Smith -> Person  
Mr. Ahmed-> Person  
aka -> Alias  
was seen -> location  
Church St. -> Address  
Magnolia Ave. -> Address  
Nov 13<sup>th</sup> -> Date

**Events/Sentiment Extraction**

Mr. Ahmed in database  
wanted for questioning  
**Suspect**  
-> send agent to this  
location

**Combined with structured data**

70's                      80's                      90's                      Now

© 2007 SPSS Inc. 12

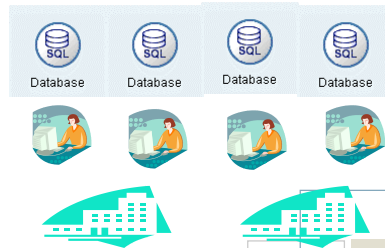


## > Predictive Asset Management



- In single project there is the potential to create a large number of models and versions of models:
  - different predictions
  - different algorithms
  - different settings
  - different training samples.

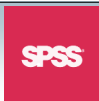
- ✘ # different data sets
- ✘ # different users
- ✘ # different locations.



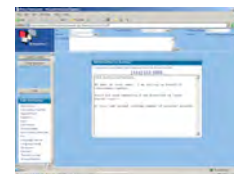
© 2007 SPSS Inc.

13

## > Feedback Collection



Data Warehouse Solutions



© 2007 SPSS Inc.

14



SPSS

### ➤ Predictive Modeling Scenarios by User

SPSS

### ➤ Predictive Tax Compliance

```
graph LR; Register([Register]) --> Assess([Assess]); Assess --> Collect([Collect]);
```

**Fraud Detection**

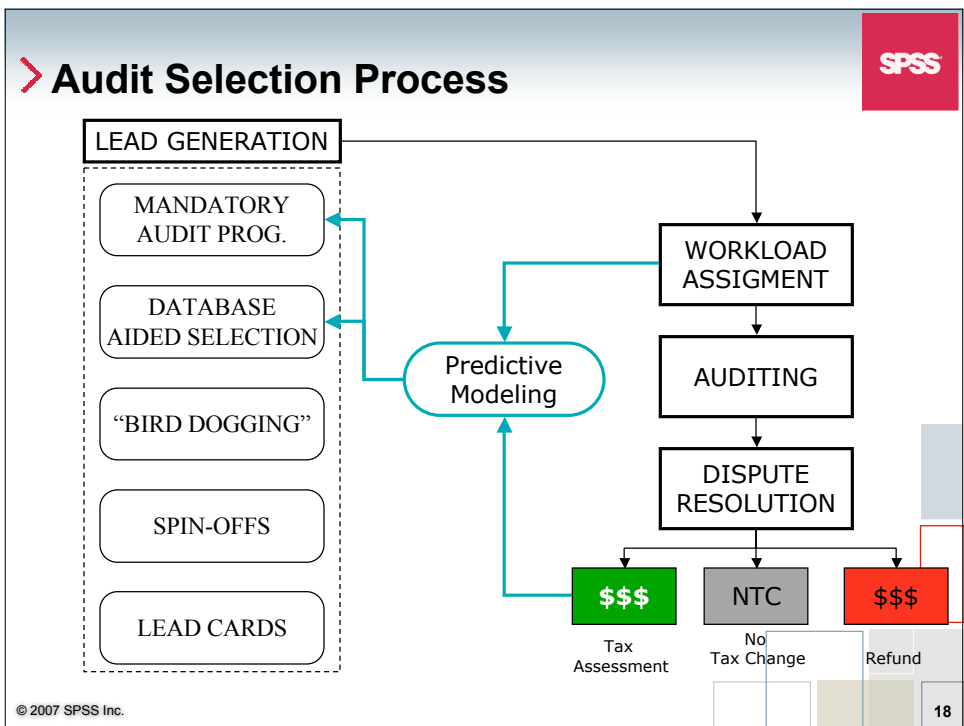
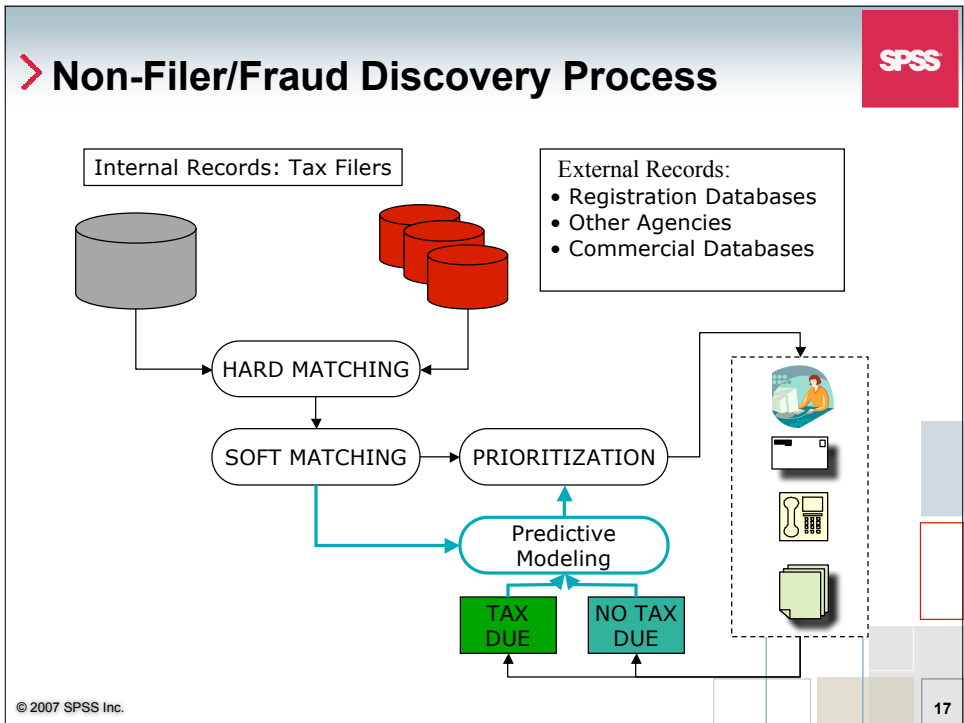
- Non-Filer Discovery**
  - ID Fraud
  - Prioritization Models
- Audit Selection**
  - Filing Fraud
- Tax Collection**
  - Treatment
  - Activities

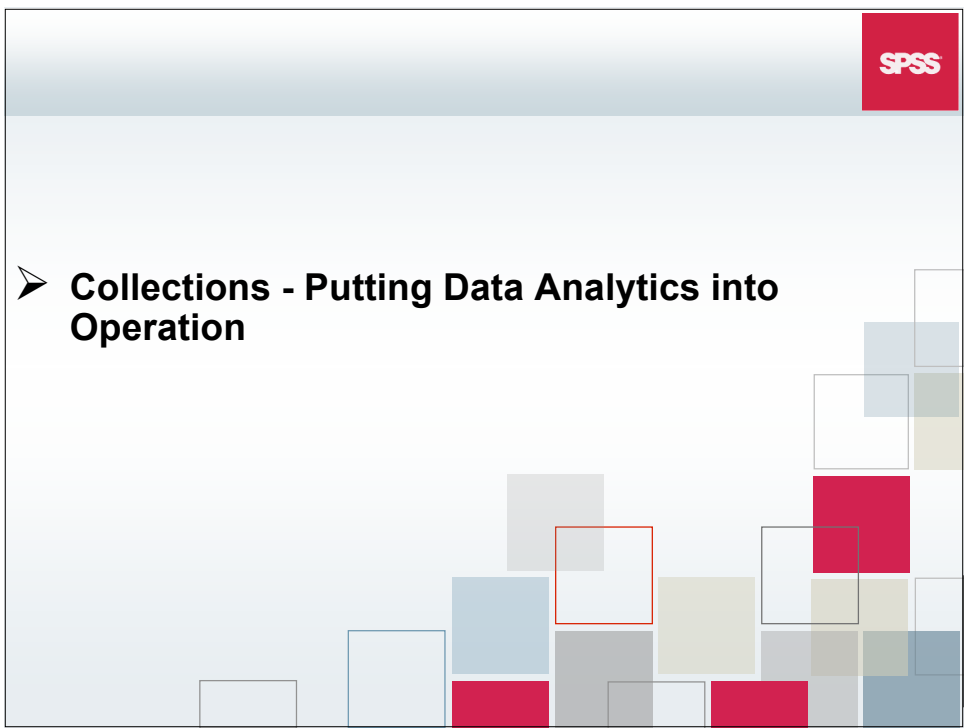
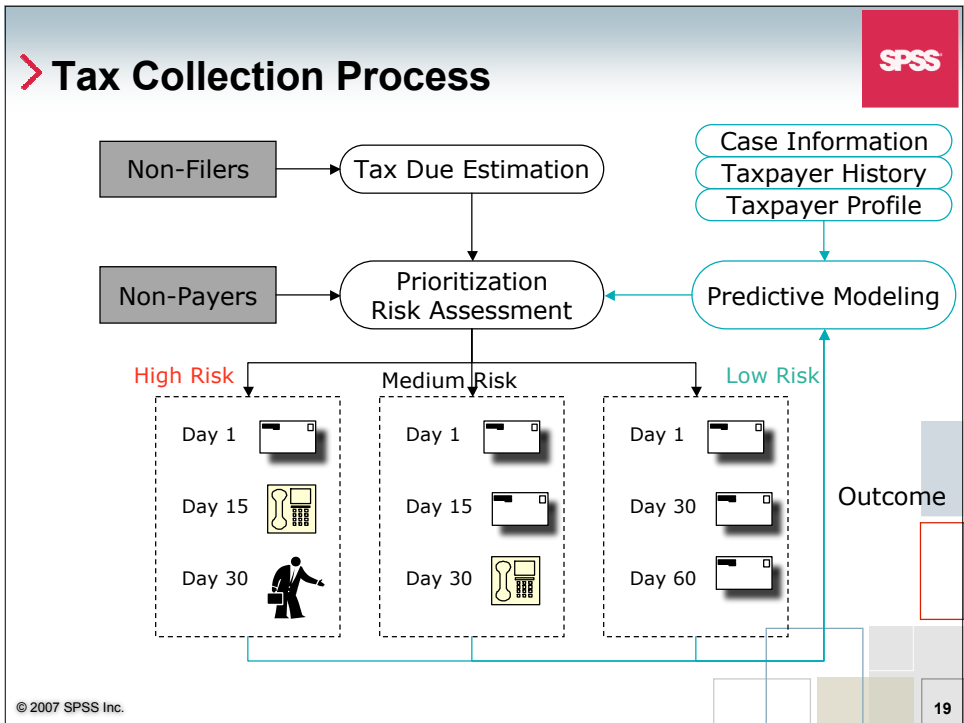
**DATA MINING & PREDICTIVE ANALYTICS TOOLS**

**DATA WAREHOUSE**

*Right work to the right resources at the right time*

© 2007 SPSS Inc. 16





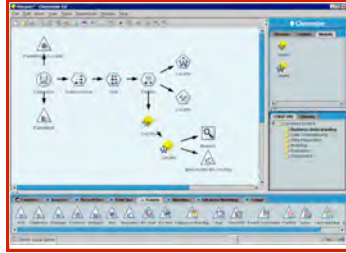


## > Collections Scenario – Putting Data Analytics into Operation



Clementine – Data Mining Workbench

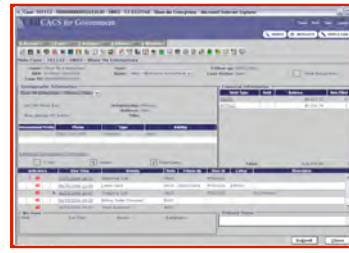
CGI CACS for Government



ROI



FBK



Data Warehouse Solutions



## > Putting Data Analytics into Operation CGI's CACS-G Collection Application



Case: 101112 - 00000000005612020 - HR03 - 12-8337768 - Show Me Enterprises - Microsoft Internet Explorer

CGI CACS for Government

SEARCH WORKLISTS OPEN A CASE

Access Case Actions Admin Windows

Main Case - 101112 - HR03 - Show Me Enterprises

Name: Show Me Enterprises  
 DBA: Systemic Unlimited  
 Case ID: UDP00005612020

User: [dropdown]  
 State: [dropdown]

Follow-up: 08/01/2006  
 Case Status: Open

Demographic Information

Address: 400 NW Mosk Ave, Blue Springs MO 64014

Relationship: Primary  
 Address: Main  
 Title:

International Profile

Phone	Type	Validity
(314) 224-0000	Business	Valid

Financial Information

Debt Type	Hold	Balance	Non Filed
SALES		\$6,423.52	0
WITHLD		\$8,316.34	2
Total:		\$14,739.86	

Additional Demographic Information

Indicators	Date Time	Activity	State	Follow Up	User ID	Letter	Descriptor
	07/03/2006 08:47	Telephone Call	HR03		PF10G0D		
	07/10/2006 12:41	Letter Sent	HR15	08/01/2006	PF10G0D	DEMO3	
	08/28/2006 14:47	Outgoing Call	HR15		PF10G0D	No Promise	
	05/15/2006 09:02	Billing State Changed	BL02				
	06/27/2006 10:00	State Accounted	HR01				

We Have

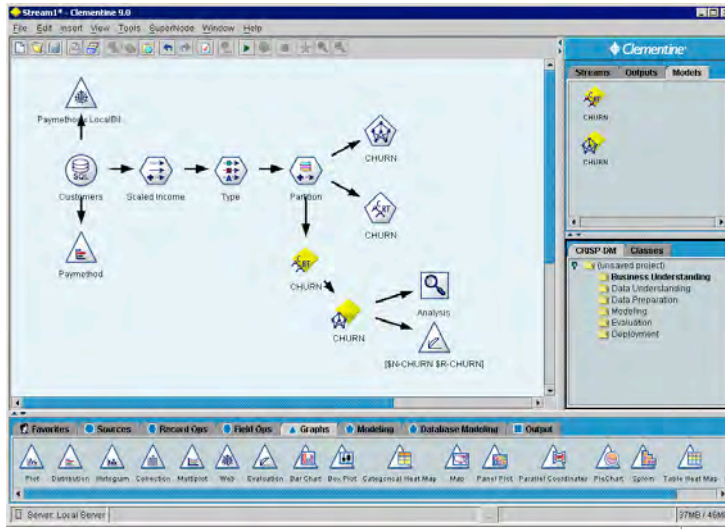
- Pay Plan
- Assets
- Bankruptcy

Submit Close





## > Data Mining Workbench



© 2007 SPSS Inc.

23

## ➤ Case Study in Audit Selection

### Evaluating Model Performance





## > Case Study in Audit Selection



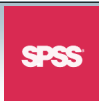
- Build models to predict different outcomes.
  - Positive Adjustment (Y/N).
  - DPH Return (Group Membership).
  - Actual \$\$ Adjustment.
- Historical Cases selected for model build
  - Cases with Prior audit – prior audit and organizational data.
  - All Cases – organizational data only.
- Deployment
  - For each outcome combine predictions for those with and without previous audit data .
  - For each outcome predict using organizational data only.

© 2007 SPSS Inc.



25

## > Evaluating Predictive Modeling



	M1	M2	M3	M4	M5	M6	Average
Audit Assessments based on Selection	144,764,062	168,856,823	143,034,137	143,285,659	212,808,908	225,188,773	1,037,938,362
Average per Audit	144,764	168,857	143,034	143,286	212,809	225,189	172,990
As a Percentage of our Best at 177,000,000	82%	95%	81%	81%	120%	127%	98%
As a Percentage of our Average at 144,000,000	101%	117%	99%	100%	148%	156%	120%
As a Percentage of our Worst at 104,000,000	139%	162%	138%	138%	205%	217%	166%

Results Based on Common Picks:	Assessments	Number Picks	Average	No-Changes	%No-Changes
1 Only	36,205,284	422	85,795	298	71%
2 Common	153,596,319	1,306	117,608	752	58%
3 Common	63,307,446	287	220,583	137	48%
4 Common	103,213,769	472	218,673	242	51%
5 Common	4,376,090	23	190,265	10	43%
6 Common	11,647,096	17	685,123	0	0%

© 2007 SPSS Inc.



26



**SPSS**

## > Optimizing Case Management with Predictive Tax Compliance

Come see us in the Exhibit Hall

**SPSS**

**Benjamin Chard**  
Senior Solution Engineer  
[bchard@spss.com](mailto:bchard@spss.com)

**Lisa Barber**  
Account Executive  
[lbarber@spss.com](mailto:lbarber@spss.com)

**CGI**

**Ted London**  
Executive Consultant  
Tax and Revenue Management  
[ted.london@cgi.com](mailto:ted.london@cgi.com)

**SPSS**

© 2007 SPSS Inc.

28



## > ROI – Return on Information



### Deploy Scores

- Build the model then create and save the scores to a data table which Case Management can access.

	Tau ID	Prob of Closing <6months
1	100001.000	0.993
2	100002.000	0.987
3	100003.000	0.917
4	100004.000	0.983
5	100005.000	0.993
6	100006.000	0.983
7	100007.000	0.993
8	100008.000	0.993
9	100009.000	0.917
10	100010.000	0.993
11	100011.000	0.993
12	100012.000	0.993
13	100013.000	0.936
14	100014.000	0.993
15	100015.000	0.993
16	100016.000	0.993
17	100017.000	0.917
18	100018.000	0.936
19	100019.000	0.936
20	100020.000	0.993
21	100021.000	0.993
22	100022.000	0.993
23	100023.000	0.917
24	100024.000	0.993
25	100025.000	0.936
26	100026.000	0.993
27	100027.000	0.993
28	100028.000	0.993

### Deploy Model

- Build the model and then save the model and scoring process to files that can be executed real-time.

