NEW YORK'S FRAUD ANALYSIS AND SELECTION TEAM

Leveraging Business Analytics to Stop Refund Fraud

In late 2011, the New York State Department of Taxation and Finance undertook a new approach to fraud recognition and detection. By marrying state-of-the-art technology with well-trained human resources, we significantly expanded our ability to detect and stop premeditated tax refund fraud and identify theft.

On the technology side, the primary tool was the Department's business analytics system that dramatically improved audit case selection prior to refunds being issued. For personnel, a new Fraud Analysis and Selection Team (FAST) leveraged that system to stop more than $150 million in fraudulent refunds in just two years.

# EVOLUTION OF BUSINESS ANALYTICS

Phase One: Refunds

The seeds of this new initiative were planted more than a decade ago, sparked by a need for a more data-driven, accurate audit selection process - a way to root out questionable tax returns and refund requests. Historically, erroneous – and even fraudulent – refund requests were typically detected and acted on only after refund checks had been sent. On top of being time-consuming and costly, efforts to recover those refunds often proved fruitless.

How could the Department catch and rectify such refunds before they went out the door? The answer lay in the sophisticated business analytics the Department had already used for audit purposes. Previously, the analytics had only been used well after returns were processed and refunds had been issued.

The use of predictive intelligence to dynamically determine when to process a refund request and when to set it aside for further analysis came to be known as the Case Identification Selection System (CISS).

Beginning in 2003, CISS fundamentally repositioned the role of business analytics in the tax return process. Because CISS embedded analytics directly into the mainstream return process, the Department was able to automatically review each of the 10 million personal income tax refunds it receives annually.

Moreover, in contrast to the relatively fixed algorithms drawn from historical data, CISS is designed to both continually improve its accuracy and stay ahead of newly emerging fraud threats. CISS is saving New York taxpayers $370 million annually by stopping erroneous refunds.
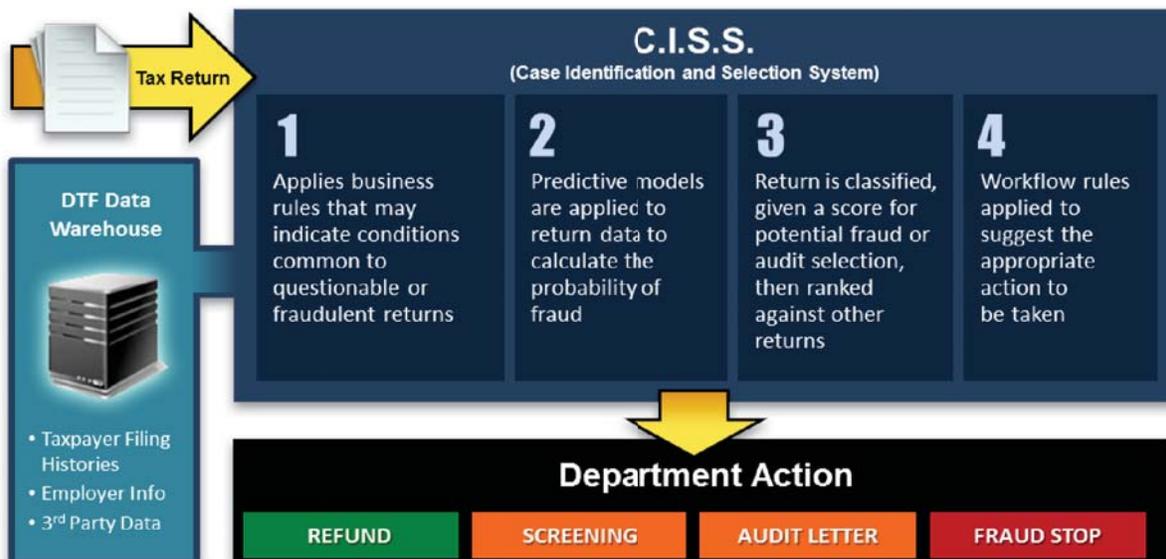
Phase 2: Withholding

In 2008, the Department extended CISS to review withholding claimed on personal income tax returns. In this case, the business analytics corroborates wages and withholding claimed by taxpayers to that reported to us by employers. Various automated workflow actions are applied to the findings.

In addition to identifying and saving more than $45 million annually, the second phase of CISS is significant for the Department when it comes to recognizing, and subsequently addressing, identity theft.

Phase 3: Collections

Results are contagious. Once CISS proved highly effective in both the refund and withholding functions, it was a natural extension to bring it to the collections process.

Delinquent tax collection is by most measures an inherently inefficient process that is reliant on letters, phone calls and the shoe-leather visits of debt collectors. Any revenue department's challenge is to determine how to allocate its limited time and personnel resources in ways that produce the most recovered revenue.

CISS Collections employs business analytics to compare each collections case with profiles of past similar cases. The result is predictive models that indicate which cases should be pursued and through which means.

For example, CISS might analyze a specific segment of taxpayers with similar characteristics – perhaps dollar range of debt, no past payment plans, and no response to prior collections notices. Then, based on past collections actions – letters, warrants, levies, and field visits – the analytics might determine that field visits will result in the highest dollar collection at the least cost.

Phase 3 of CISS is generating an additional $80 million in collections annually.

FAST – FRAUD ANALYSIS AND SELECTION TEAM

While CISS was able to review individual returns and develop predictive models, it was not a system intended to or identify patterns among multiple returns. Enter the Fraud Analysis and Selection Team (FAST) – two full-time audit supervisors and seven staffers who contribute to the team.

FAST analyzes income tax returns prior to processing to target large-scale refund fraud schemes. Initially, CISS provides selection and scoring data for every return, which enables the team to narrow its focus.

In addition to CISS, the team leverages various other technologies to identify patterns that suggest fraud among millions of returns. Initially, Statistical Analysis System (SAS) reports are run weekly to garner the necessary data for further analysis. During peak season, the reports can include up to one million returns.

FAST members access the weekly reports to begin their fraud queries. Excel pivot tables are used to identify commonalities among multiple returns, including:

- same paid preparer
- common Internet protocol (IP) or email addresses
- redundant - or non-existent - mailing or physical addresses
- common bank accounts for direct deposit
- similar refund amounts
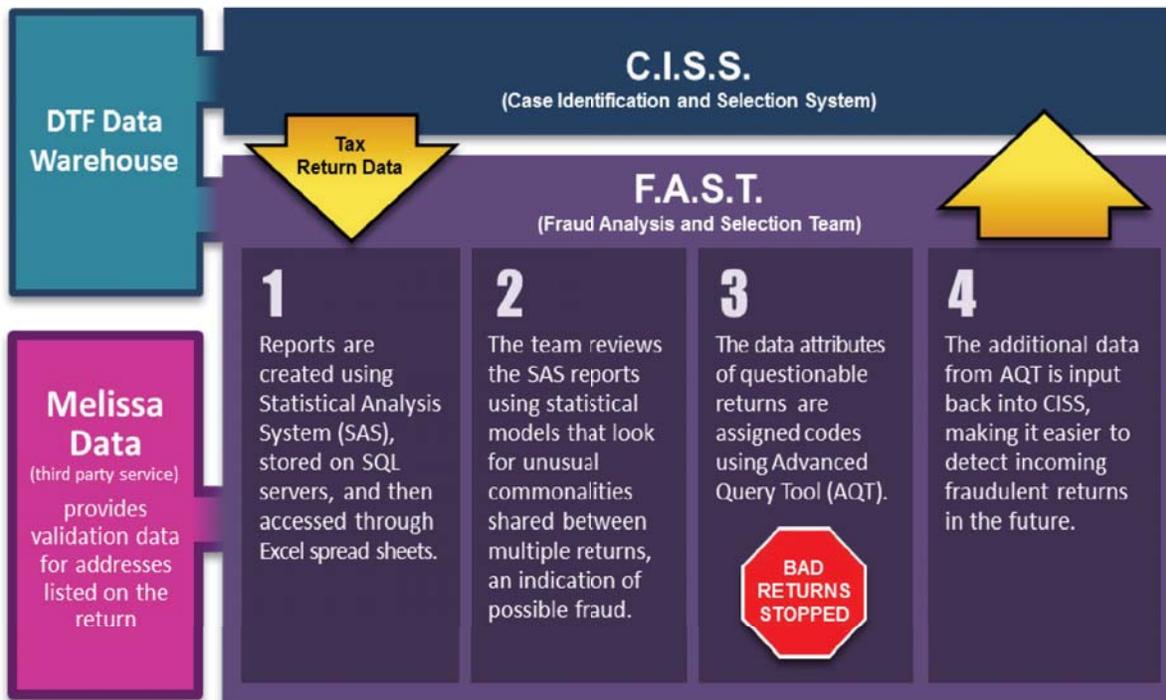- repetitive credit claims
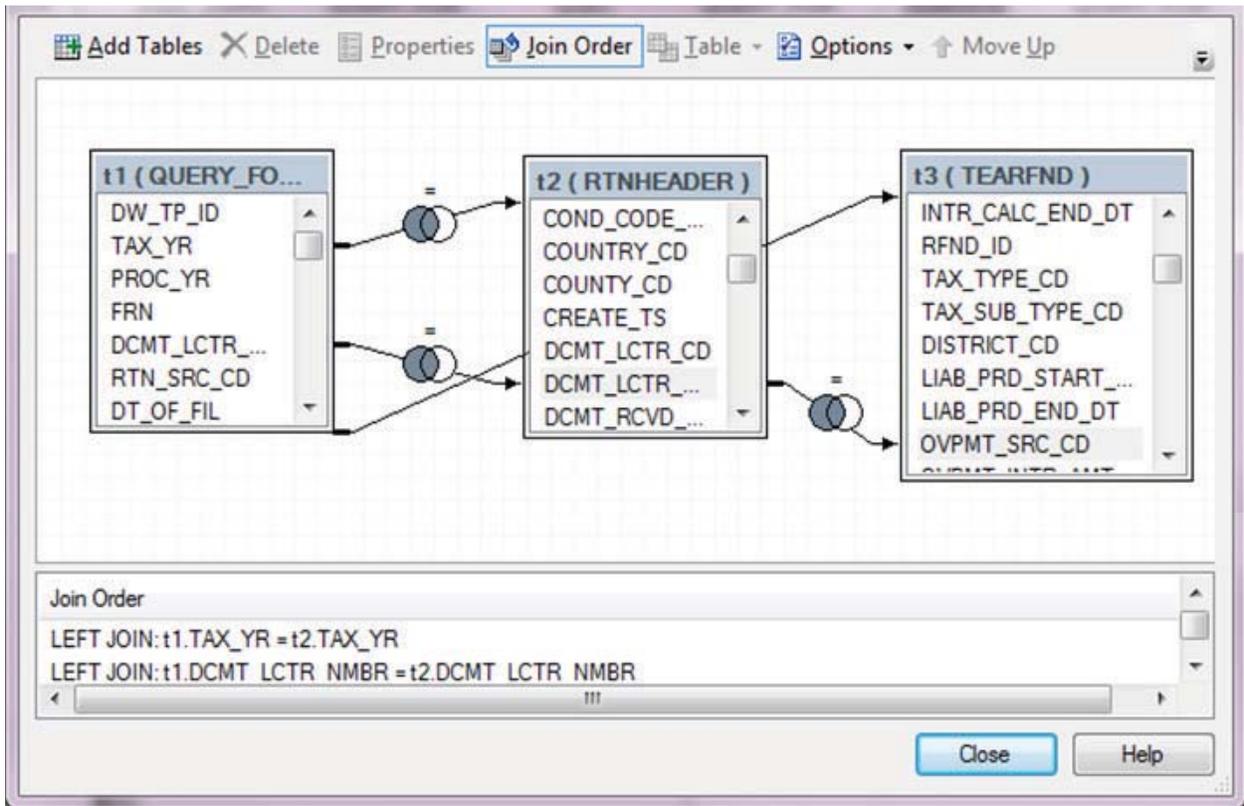
The screenshot above shows tables joined in SAS. In this case the CISS Table is joined with the Personal Income Tax Return Table and the Refund Table. The resuling data indicates the outcome of each return for that week, including whether a refund was stopped by CISS.

The example below of a SAS report is a likely case of fraud. More than 120 returns were filed from nine similar IP addresses – all from the Fort Lauderdale area. All of the returns were asking for refunds greater than $150 and we were unable match any of them to employer data. FAST coded the IP addresses to stop any refunds from being issued on associated returns.

IP Address	City	Region	Country	num returns	num SINGLE	% SINGLE	num direct deposit	% DIR DEP	num refunds > 150	% rfnd gt 150	NUM_NOT_IN_WRS	% not in WRS
069.112.214.069	Patchogue	NY	US	19	16	84	17	89	19	100	0	0
072.080.125.031	Jamaica	NY	US	11	6	54	11	100	11	100	0	0
098.113.225.032	Staten Island	NY	US	28	15	53	25	89	26	92	1	3
070.193.197.112			US	25	25	100	25	100	25	100	25	100
070.193.197.123			US	10	10	100	10	100	10	100	10	100
070.193.198.078			US	12	12	100	12	100	12	100	12	100
070.193.198.198			US	15	15	100	15	100	15	100	15	100
070.193.199.178			US	10	10	100	10	100	10	100	9	90
070.193.206.166			US	10	10	100	10	100	10	100	10	100
068.039.234.079	Somerset	NJ	US	23	21	91	23	100	23	100	0	0
050.140.202.099	Hollywood	FL	US	36	36	100	36	100	36	100	36	100
064.013.126.181	Anchorage	AK	US	93	91	97	93	100	93	100	93	100
100.001.094.175	Jersey City	NJ	US	14	6	42	14	100	13	92	0	0
024.046.010.121	Bronx	NY	US	11	5	45	11	100	11	100	0	0
069.180.105.095	Fort Lauderdale	FL	US	13	13	100	13	100	13	100	13	100
070.193.192.202	Fort Lauderdale	FL	US	11	11	100	11	100	11	100	11	100
070.193.193.110	Fort Lauderdale	FL	US	27	27	100	27	100	27	100	27	100
070.193.193.127	Fort Lauderdale	FL	US	19	19	100	19	100	19	100	19	100
070.193.193.164	Fort Lauderdale	FL	US	15	15	100	15	100	15	100	15	100
070.193.194.041	Fort Lauderdale	FL	US	13	13	100	13	100	13	100	13	100
070.193.195.097	Fort Lauderdale	FL	US	28	28	100	28	100	28	100	28	100
069.115.129.243	Wappingers Falls	NY	US	11	9	81	11	100	10	90	0	0

3

Once a scheme is discovered, Advanced Query Tool (AQT) is used to code preparers, IP addresses, bank routing numbers and individual social security numbers. This will either question the return with an audit letter or fully stop returns that are clearly fraudulent.

In addition, CISS is brought back into the process to look for the returns with matching codes and to take appropriate action. CISS then stops the scheme from that point forward – for that processing year, and all years in the future.

C.I.S.S.
(Case Identification and Selection System)

DTF Data Warehouse

Tax Return Data

F.A.S.T.
(Fraud Analysis and Selection Team)

Melissa Data
(third party service)

provides validation data for addresses listed on the return

1 Reports are created using Statistical Analysis System (SAS), stored on SQL servers, and then accessed through Excel spread sheets.

2 The team reviews the SAS reports using statistical models that look for unusual commonalities shared between multiple returns, an indication of possible fraud.

3 The data attributes of questionable returns are assigned codes using Advanced Query Tool (AQT).

BAD RETURNS STOPPED

4 The additional data from AQT is input back into CISS, making it easier to detect incoming fraudulent returns in the future.

Case Study: Tax Preparer Pleads Guilty

Approximately 70% of New York taxpayers use a tax preparer to complete and file their returns. Here's how FAST helped investigators shut down a fraudulent, Bronx-based tax preparer who falsely claimed more than $7 million in refunds.

During review of SAS reports, FAST noticed that 80% of a certain preparer's returns claimed a refund of exactly $263, comprised of the College Tuition Credit and the NYC School Tax Credit. The College Tuition Credit claims were verified as fraudulent because they all listed the same college, without providing an Employer Identification Number for the college. The automated audit selection rules were adjusted so that future claims from this preparer would be reviewed.

Shortly thereafter, the Department began to receive inquiries regarding this preparer, associated with a firm named E & M Multiservices. The Bronx District Attorney's office called the Tax Department to inquire about rumors of money that was being issued by this preparer to senior citizens from "President Obama or the estates of Whitney Houston and Steve Jobs". A news outlet called for information about lines running around the block of the 99¢ store where the preparer ran his office.

FAST ran reports for in-depth reviews of the preparer's filing activities. It was apparent that he was preparing returns in the names of seniors claiming fictitious business income. The team gathered information for our Criminal Investigations Division; that unit worked with other law enforcement authorities to create a case against the preparer. Eventually the investigation and prosecution was handed over to federal authorities.

In August, the tax preparer pled guilty to one count of subscribing to a false and fraudulent tax return for himself, one count of aiding and assisting in the preparation of a false tax return for a relative and one count of wire fraud. He faces a maximum sentence of 26 years in prison and is scheduled to be sentenced later this year.

Case Study: Home Attendant Business Losses

New York, like many states, offers a choice of receiving refunds via debit cards, direct deposit or paper check. A 2012 FAST report on refund debit cards over $5,000 brought back results of several returns claiming large business losses to bring incomes with higher wages into the "sweet spot" for the Earned Income Tax Credit. Many returns with Eastern European names from the Brighton Beach area of Brooklyn were filing similar returns – claiming real wages and significant business losses as home health aides, home attendants or street vendors.

FAST went to work to find more returns with similar characteristics. Audit cases were created to request documentation to support the business losses claimed. Documentation was not provided. Refunds were denied and bills issued for those who owed tax.

The findings of this case led to identification of several other schemes related to business losses. Thus far, the Department collected $5 million from more than $9.5 million in assessments issued. In turn, the CISS audit selection rules have been adjusted to put similar refund requests on automatic hold, while a letter is sent to the taxpayer to request verification of the business loss.

RETURN ON INVESTMENT

Since its inception just over two years ago, FAST has saved New York taxpayers more than $150 million - a significant portion of the overall CISS savings.

In contrast, FAST staff costs total $600,000 annually. Even when factoring in the initial cost of CISS ($3.8 million) and additional software ($7,000) the return on investment is tremendous.

Total Return $151M

■	2014*	$14M
■	2013	$73M
■	2012	$64M

Total Investment $5.15M

■	Infrastructure	$3.8M
■	Labor	$1.35M

* year to date

CISS NEXT GENERATION

CISS and FAST continuously improve and update the universe of cases handled. Where to in the future? The next frontier will be sales tax, the ability to analyze returns as they're filed and critically draw from third-party data to review them as they're received.

In addition, the Department is establishing the Tax Analytics Solution Center (TASC) to incorporate into CISS the ability to recognize the filing patterns currently identified by FAST. The result will be an automated capacity to stop multiple refunds based on commonalities.

Focusing on fraud, TASC will also provide predictive models based on patterns, historical data, and third party data, as well as give FAST new tools (including SPSS Modeler) to evaluate returns faster and in real-time.

CONCLUSION

It's one thing to analyze questionable tax refunds after they've been sent out, when the only option left is to pursue them – often to no avail. It's another to detect questionable refunds before they are issued, preventing the waste of resources and the loss of critical tax revenue.

The New York State Department of Taxation and Finance has creatively and innovatively integrated predictive analytics and auditing talent directly into its processing stream, thus preemptively identifying questionable tax returns and tax refund schemes, while optimizing the Department's approach to collecting delinquent taxes.

The results: a dramatic reduction in the dollar amount of questionable refunds paid, administrative dollar savings associated with the abandonment of "pay and chase" policies, improvements in delinquent tax collections, and greater inter-departmental collaboration on cases, especially between Audit and the Criminal Investigations Division.

We look forward to providing more information or addressing any questions you may have.