



Moving to the Future Data Mining & Data Modeling

Presented by:
New York State Office of the State Comptroller
Division of State Government Accountability
Medical Claims Unit



Agenda Topics

- MCU & Medicaid
- Data Tools
- Data Analysis
- Data Mining
- Using Results



Medical Claims Unit

- 15 staff members:
- Audit payments to Medicaid providers
 - Weekly review of Medicaid payments
 - Conduct reviews of program operations & Medicaid providers



New York State Medicaid Program

- \$46 Billion in Payments per Year
- Over 4 Million Recipients
- Over 100,000 Providers
- An Average of 1 Million Claims Daily



Data Tools

- Excel
- ACL
- **Clementine**
- Arc GIS



Data Analysis

- Snapshot of a business at a given point in time
- Identifies extreme outliers
- Focuses audit efforts toward risk
- Identifies low lying fruit



Typical Analysis Steps

- Duplicates
- Highest Paid
- Trending
- Distance Analysis
- Criteria Driven Analysis

Dr. Mayell

Value	Proportion	Count	DOB	Record_Count
D2391	30.68	4914	2004-06-30	9144.000
D2390	18.72	2400	2003-10-30	2418.000
D2392	6.47	847	2005-03-11	3984.000
D0730	5.51	687	2003-10-17	2782.000
D2393	4.71	616	2004-08-09	2624.000
D2140	4.61	603	2005-02-24	3088.000
D1110	3.81	496	2003-04-02	2696.000
D0640	3.71	485	2004-09-30	2734.000
D2294	3.11	406	2003-10-03	2934.000
D0200	2.7	351	2006-09-12	2840.000
D0272	2.25	294	2004-07-08	2956.000
D0200	1.92	251	2006-09-03	2630.000
D2181	1.45	190	2006-06-03	3303.000
D0600	1.14	150	2004-03-31	2624.000
D2200	1.22	160	2003-06-03	2366.000
D2160	1.21	158	2004-01-08	2224.000
D1200	1.18	152	2004-08-10	2586.000
D1200	0.78	100	2006-10-13	2400.000
			2005-04-20	2600.000
			2006-05-15	3212.000
				29



Audit Results

- \$896,000 in Questionable Payments
- \$125,000 Reviewed
- 100% Percent Disallowed

<http://osc.state.ny.us/audits/allaudits/093007/07s3.htm>



Dr. Zukor

- Billing Spikes (Trending)
- Filling all 32 teeth
- Extracting all 32 teeth
- Downstate Doc, Upstate Patients



Dr. Zukor

- \$1.7M in Questionable Payments
- \$307,000 Reviewed
- 100% Percent Disallowed

<http://osc.state.ny.us/audits/allaudits/093008/08s67.htm>



What Are We Missing?

- The not so obvious
- The mathematical explanation
- The "bad" providers hiding in the norm



Data Mining

- Unknown useful information in your data
- Statistical and logical analysis of large sets of transaction data
- Pattern recognition and machine learning



Using What We Know

- Find useful information in our data
- Use sophisticated software to do the work
- Find unknown “bad” billers who bill similar to Mayell and Zukor.



What Are We Going to Do?

- Create new data analysis variables based on known findings
- Have software algorithms do all the mining work
- Investigate data mining outliers based on recognized pattern(s)



What We Did

- Created MCU defined variables
- Brought summary data to the claim level
- Built Linear Regression Model to predict payment based on relevant variables
- Calculated overpayments based on the linear regression model's payment equation



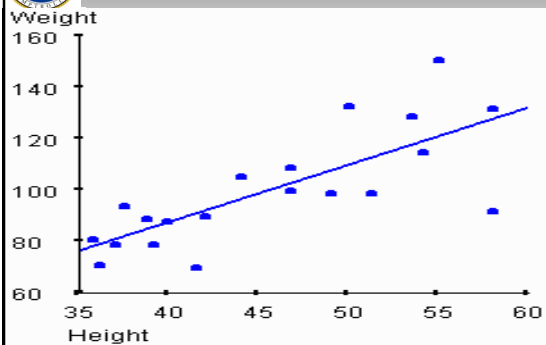
Linear Regression Model

$$y = mx + b$$

- A statistical technique that correlates the change in one variable to other variable(s).
- x_1 through x_n predicts y



Linear Regression Example





Initial Results

- 38 providers identified as potential outliers
- Identification of questionable billing patterns
- Confirmation of other known bad providers not characterized
- 17 focused investigations

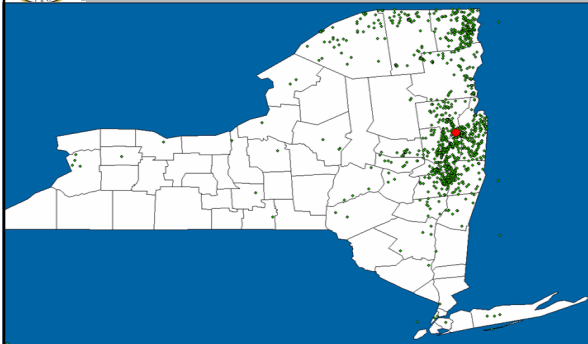


Findings

- Billings on every day of the calendar year
- Billing for long distance recipients
- Providers sanctioned and/or no longer allowed to bill Medicaid



Mapping Information





Moving Forward

- Determine the model's strengths
- Add/Subtract variables to the model based on the results of investigations
- Use the model's principles to build new models for other provider types
- Explore other modeling techniques (neural networks, clusters, associations, sequencing, decision trees)



Contact Info

NYS OSC Medical Claims Unit
One CSC Way
Rensselaer, NY 12144
(518) 474-3271
dtowle@osc.state.ny.us
