Using Computer-Assisted Auditing Techniques to Detect Fraud
Using CAATs to Find Fraud

- What are CAATs?
- Data Analysis Methodology
- What is Data Mining?
- Fraud Detection
- Data Analysis Software & Techniques
- Examples of Fraud
- A Generic Approach
- Benford’s Law
- Financial Crime Investigator
- Case Studies
What Are CAATs?

– Computer-based tools that permit auditors to increase their personal productivity as well as that of the audit function. [CAATTs & Other BEASTS for Auditors, by David G. Coderre; 1998, Global Audit Publications]

– Provides, at a minimum, the following benefits:
  – Gain insight into the business and operations
  – Visibility into the company’s control (failure, operations metrics, improvements)
  – Benchmarking across business units, competitors, etc.
  – Testing efficiencies – increase test coverage and assurance
  – Regulatory compliance, fraud or litigation analysis
  – Saves time (sometimes even in the first year)
What Are CAATs?

- *The data tells the story!* Enables you to quantify the financial impact of business decisions, accounting practices, and internal controls
- Also known as “Data Analysis”
- The power of CAATs: The Georgia Department of Redundancy Department has 135,000 employees. How long would it take auditors to manually search payroll records to identify duplicate payments by finding duplicate social security numbers?
Data Analysis Methodology

1. Define
   1.1 Define objectives of analysis
   1.2 Gain an understanding of business/technical environment
   1.3 Define data requirements

2. Validate
   2.1 Request and receive data
   2.2 Validate control totals
   2.3 Perform data quality assessment
   2.4 Gain understanding of data

3. Execute
   3.1 Execute audit steps
   3.2 Identify discrepancies
   3.3 Discuss discrepancies with stakeholders and validate errors
   3.3 Assess impact on objectives

4. Data Retention
   4.1 Document process to reproduce data

Document
Types of CAATs

- Word processing
- Spreadsheet
- Database
- Statistical sampling
- Data mining
- Real time testing programs
- Integrated audit software
- Data analysis
- Artificial intelligence/expert systems
Types of CAATs Most Important for Fraud Detection

- Word processing
- Spreadsheet
- Database
- Statistical sampling
- Data mining
- Real time testing programs
- Integrated audit software
- Data analysis
- Artificial intelligence/expert systems
Using Data Analysis Software

1. **Test Objectives**
2. **Identify Data**
3. **Request Test Data**
4. **Run Test CAATS**
5. **Test Results**
   - **Good**
     - **Request Full Data**
     - **Run CAATS**
   - **Poor**
     - **Refine Data Request**

**CAAT Reports**
What is Data Mining?

◆ The process of discovering meaningful new correlations, patterns, and trends by sifting through large amounts of data stored in repositories, using pattern recognition technologies as well as statistical and pattern recognition techniques [Gartner Group Interactive: http://gartner6.gartnerweb.com]

◆ Most often used (up until recently) in marketing and customer analysis
Data Mining in Crime Solving

- Software compiles facts, attributes, and characteristics about various types of crimes
- Helps investigators identify crimes with common (or similar) attributes/characteristics
- Linking evidence from similar crimes can lead to identification of perpetrator(s)
- Detective Toolkit (Violent/Serial Crime)
- Fraud Investigator (Insurance fraud)
- Similarity Search Engine (compares any databases)
Data Mining Crime Solving Example

- Insurance companies compile data on claims: incident descriptions, claimants, witnesses, other individuals involved, time of day, location, etc.
- Data mining software identified cases where the same individual was involved in several claims, sometimes as witness, sometimes as passenger, sometimes as driver.
- Further comparisons and investigation lead to identification of hundreds of fraudulent claims.
Data Mining Crime Solving Example

- A series of murders occurred with remarkable regularity (weekly) in a small town in Maine.
- Police used data mining techniques to track the similarities and common characteristics of all of these crimes.
- They found that every crime had a single common characteristic; one person was involved in some way with every one of the murders.
- Although they were unable to prove that this person had committed the crime, Cabot Cove declared Jessica Fletcher a public menace, banned her from the town, and the murders stopped.
Getting Data

Never! Never! Never! Never! Give Up!

Winston Churchill
Fraud Detection
Fraud Detection

- Think “outside the box” … “one plus one equals two” is not always [Fraud Examination in the Classroom, by Mary-Jo Kranacher, May / June 2005, Fraud Magazine]

- Batman once said, “If only they would use their genius for good instead of evil!”
What is Fraud?

**Employee Fraud**
Activity to benefit himself and affect the company

**Management Fraud**
Activity to benefit the company
Fraud Tests

- Asset misappropriation
- Fraudulent Statements
- Corruption

Audit Program
FRAUD Tests
Fraud Detection Plan

- Hypothesis Testing
  - Develop a fraud hypothesis
  - Obtain data
  - Design CAAT tests
  - Analyze results to determine if there is support for fraud hypothesis
How Can We Use Data Mining to Find Financial Statement Fraud?

- Compile databases of key ratios, industry characteristics, and other attributes (risk factors) of discovered financial statement frauds
- Use data mining techniques to calculate coefficients of correlation between known financial statement fraud schemes and the organization you are planning to audit
- Results imply the degree of audit risk (and have corresponding implications about audit fees)
- Better still, results will pinpoint the areas within the financial statements needing the greatest audit attention
- In effect, a more sophisticated type of analytical procedure than we have done traditionally
Data Analysis Software

- Useful for identifying misappropriation of assets and fraudulent financial reporting
- Allows limitless number of analytical relationships to be assessed
  - within large databases
  - comparing large databases
- Identifies anomalies
- Further (human) investigation is almost always needed
Data Analysis Software

- Access and Excel
- Interactive Data Extraction and Analysis (IDEA)
- Audit Command Language (ACL)
- Windows based and user friendly
- Require creativity and imagination
- Supplements – but does not replace – intelligent audit work
Data Analysis Techniques

- Filters
- Sorts
- Statistics
- Gaps
- Duplicates
- Aging
- Confirmations
- Samples

- Classification
- Summarization
- Stratification
- Join and Define Relationships
- Trend Analysis
- Regression Analysis
- Parallel Simulation
- Digital Analysis
## Data Analysis Tools

<table>
<thead>
<tr>
<th>Tools</th>
<th>Capacity</th>
<th>Ease of use</th>
<th>Analytic Capabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microsoft Excel</td>
<td>• 65,536 rows by 256 columns</td>
<td>• Standard, easy to use office application</td>
<td>• Data analysis toolkit • Built-in functions</td>
</tr>
<tr>
<td>Microsoft Access</td>
<td>• 2 GB database • 255 fields (columns)</td>
<td>• Training is required</td>
<td>• Built in functions • Great for joining tables</td>
</tr>
<tr>
<td>ACL</td>
<td>• Unlimited</td>
<td>• Requires basic training • Menu based</td>
<td>• Complete set of preprogrammed analysis</td>
</tr>
<tr>
<td>Monarch</td>
<td>• 1,000,000 input pages</td>
<td>• Training is required</td>
<td>• Provides basic analytic capabilities</td>
</tr>
<tr>
<td>Microsoft SQL Server</td>
<td>• 1,048,516 terabytes • 1,024 columns</td>
<td>• Advanced training required</td>
<td>• Built in functions • Great for joining tables</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Ease of use</strong></th>
<th><strong>Excellent</strong></th>
<th><strong>Good</strong></th>
<th><strong>Satisfactory</strong></th>
<th><strong>Poor</strong></th>
<th><strong>Very poor</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Microsoft Excel</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Microsoft Access</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ACL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Monarch</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Microsoft SQL Server</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CARTA

Composite Application, Right-Time Architecture (CARTA)
Examples of Fraud

- **Fraud Analysis** – A simple analysis of data such as payroll, employee, vendor, accounts payable, accounts receivable, and much more, can help determine if fraud is occurring

- **Payroll Fraud**
  - Duplicates (i.e. payees on same date, same or similar names, direct deposit account numbers)
  - Paychecks being created for employees that have no time and attendance, no expenses, no vacation, little or scarce personnel records, etc
  - Wages inconsistent with job classification
  - Pay date precedes employment date
  - Terminated employees continuing to be paid
Examples of Fraud

- **Purchasing Fraud**
  - Duplicate disbursement amounts
  - Duplicate invoice numbers/dates
  - Duplicate disbursements on same date
  - Disbursement to vendor not in vendor database
  - Vendor name/address/phone # same as employee name/address/phone #
  - Invoice’s “pay to” address different from address in vendor or contract database

- **Refund Fraud**
  - Refund check “pay to” address different from address in customer database
  - Refund check amounts just below higher-level-approval-required threshold
  - Refund check “pay to” name and/or address matches to employee name and/or address
Examples of Fraud

- **Accounts Payable**
  - Producing reports of debit balances
  - Producing reports of large or old suspense items
  - Testing accumulation of payables balances
  - Producing reports of balances with no scheduled payment date
  - Producing reports of new suppliers
  - Search for unrecorded liabilities
    » These reports help identify
      - Inefficient invoice processing
      - Spend reduction opportunities
      - Inefficient purchasing organization
      - Mismanagement of cash flow
      - Inconsistent payment terms across organizations
      - Data quality issues within master files
Examples of Fraud

- **Cash Skimming**
  - Unrecorded or understated sales or receivables
  - Theft of cash receipts
  - Lapping

- **Fraudulent Disbursements**
  - Fictitious vendors
  - Billing schemes
  - Over-billing schemes
  - Pay and return scheme
  - Check kiting
  - Theft of company checks and check tampering
  - Expense reporting schemes
Examples of Fraud

- **Inventory**
  - Theft of inventory
  - False sales, write-offs and other adjustments
  - Inventory valuation schemes

- **Fixed Assets**
  - Theft of fixed assets
  - Unauthorized changes in depreciable life
  - Unsupported additional / deletions / modifications to fixed asset sub ledger
Detection Techniques: Digitizing

Use of special functions to convert names to digits, and comparison of digits for phonetic duplicates.

<table>
<thead>
<tr>
<th>Soundex Code</th>
<th>Vnum</th>
<th>Vname</th>
<th>Addr1</th>
<th>City</th>
</tr>
</thead>
<tbody>
<tr>
<td>A153</td>
<td>68</td>
<td>AVNET</td>
<td>P.O. BOX 847722</td>
<td>DALLAS</td>
</tr>
<tr>
<td>A153</td>
<td>57</td>
<td>AVNET EMG</td>
<td>P.O. BOX 847722</td>
<td>DALLAS</td>
</tr>
<tr>
<td>P626</td>
<td>251</td>
<td>PROGRAMMERS PARADISE</td>
<td>P.O. BOX 17043</td>
<td>NEWARK</td>
</tr>
<tr>
<td>P626</td>
<td>855</td>
<td>THE PROGRAMMER'S SHOP</td>
<td>P.O. BOX 17043</td>
<td>NEWARK</td>
</tr>
</tbody>
</table>
### Testing Example: Related Party Transactions

Match: Addresses, Phone Numbers, SSN-FEDID

<table>
<thead>
<tr>
<th>Vendor LOC</th>
<th>VCODE/</th>
<th>VNAME/</th>
<th>Vendor Address1/</th>
<th>Vendor City/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employee LOC</td>
<td>Employee SSN</td>
<td>Employee (fname+lname)</td>
<td>Employee Address 1</td>
<td>Employee CITY</td>
</tr>
<tr>
<td>DST2017</td>
<td>LES142</td>
<td>LESLIE SMITH WELDING</td>
<td>1420 CHERRY HILL LOOP</td>
<td>SPRINGFIELD</td>
</tr>
<tr>
<td>2017</td>
<td>333245232</td>
<td>LESLIE SMITH</td>
<td>1420 CHERRY HILL</td>
<td>SPRINGFIELD</td>
</tr>
<tr>
<td>DST0086</td>
<td>14195</td>
<td>MORRIS COMPUTER SERVICES</td>
<td>4402 CODY AVENUE</td>
<td>BELMONT</td>
</tr>
<tr>
<td>0086</td>
<td>435443344</td>
<td>GARY MORRIS</td>
<td>4402 CODY AVE</td>
<td>BELMONT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VENDORNUM</th>
<th>NAME1</th>
<th>STREET</th>
<th>c_dummy</th>
<th>Zip</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSN</td>
<td>NAME</td>
<td>ADD1</td>
<td>ADD2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0040006668</td>
<td>T ENTERPRISE</td>
<td>RT 2 BOX 8</td>
<td>MENTOR</td>
<td>56736</td>
<td>312-444-4444</td>
</tr>
<tr>
<td>232456545</td>
<td>Tischman, Richard C</td>
<td>Rt 2 Box 8</td>
<td>Mentor</td>
<td>56736-9703</td>
<td>630/4548976</td>
</tr>
</tbody>
</table>
# Testing Example: Corporate Card Misuse

## String Searches

<table>
<thead>
<tr>
<th>ACCT</th>
<th>ESTB_NAME</th>
<th>REFNUM</th>
<th>PROCESSDT</th>
<th>BILLAMT</th>
</tr>
</thead>
<tbody>
<tr>
<td>3785-016567-54000</td>
<td>GO GO TOURS</td>
<td>000000514873321</td>
<td>9/26/2002</td>
<td>2568.24</td>
</tr>
<tr>
<td>3794-143377-31000</td>
<td>TARGET 1764</td>
<td>00000001764085</td>
<td>9/27/2002</td>
<td>328.77</td>
</tr>
<tr>
<td>3785-424387-72001</td>
<td>WAL MART STORES INC</td>
<td>0000005279216</td>
<td>10/1/2002</td>
<td>310.94</td>
</tr>
<tr>
<td>3794-187357-11001</td>
<td>TARGET STORES 1122</td>
<td>0000001122112</td>
<td>9/26/2002</td>
<td>284.11</td>
</tr>
<tr>
<td>3787-226007-31002</td>
<td>THE JAMESON INN 89</td>
<td>000011441010005</td>
<td>10/12/2002</td>
<td>263.24</td>
</tr>
<tr>
<td>3785-014867-22003</td>
<td>BORDERS BOOK SHOP 23</td>
<td>000000002345556</td>
<td>9/26/2002</td>
<td>281.07</td>
</tr>
<tr>
<td>3785-297717-11001</td>
<td>BOSTON DUCK TOURS LTD</td>
<td>000000012199239</td>
<td>10/16/2002</td>
<td>261.5</td>
</tr>
<tr>
<td>3794-612827-11008</td>
<td>WAL MART STORES INC</td>
<td>000000528774995</td>
<td>9/30/2002</td>
<td>257.3</td>
</tr>
<tr>
<td>3785-016567-54000</td>
<td>WAL MART STORES INC</td>
<td>000000085987277</td>
<td>10/14/2002</td>
<td>250.73</td>
</tr>
<tr>
<td>3783-139627-03006</td>
<td>WAL MART STORES INC</td>
<td>0000009868409</td>
<td>10/19/2002</td>
<td>241.06</td>
</tr>
<tr>
<td>3785-007067-91004</td>
<td>WAL MART STORES INC</td>
<td>0000007443432</td>
<td>10/9/2002</td>
<td>232.44</td>
</tr>
<tr>
<td>3785-260367-82009</td>
<td>TARGET STORES 1186</td>
<td>0000001186074</td>
<td>10/14/2002</td>
<td>216.47</td>
</tr>
<tr>
<td>3794-000057-11003</td>
<td>THE JAMESON INN 50</td>
<td>0000441010004</td>
<td>9/26/2002</td>
<td>205.62</td>
</tr>
<tr>
<td>3785-016567-54000</td>
<td>SEARS ROEBUCK 2206</td>
<td>00000005215149</td>
<td>9/29/2002</td>
<td>204.34</td>
</tr>
<tr>
<td>3794-143377-31000</td>
<td>TARGET STORES 13</td>
<td>00000000013073</td>
<td>9/26/2002</td>
<td>195.64</td>
</tr>
<tr>
<td>3785-484307-21003</td>
<td>COSTCO 99142</td>
<td>00001913000400</td>
<td>9/28/2002</td>
<td>192.6</td>
</tr>
<tr>
<td>3794-664537-21009</td>
<td>SEARS ROEBUCK 1029</td>
<td>0000001036170</td>
<td>10/22/2002</td>
<td>183.76</td>
</tr>
<tr>
<td>3794-036757-01006</td>
<td>WAL MART STORES INC</td>
<td>00000056441284</td>
<td>10/2/2002</td>
<td>183.2</td>
</tr>
</tbody>
</table>
Testing Examples: Ghost Employees Testing

Non valid SSN:

<table>
<thead>
<tr>
<th>AH_BUSINES</th>
<th>WORK_LOCAT</th>
<th>NAME</th>
<th>SSN</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retail Operations</td>
<td>09226 Miami</td>
<td>Jones, David</td>
<td>333-24-5555 A</td>
<td></td>
</tr>
<tr>
<td>Retail Operations</td>
<td>09453 Clearwater</td>
<td>Tap, John</td>
<td>542-44-9999 A</td>
<td></td>
</tr>
<tr>
<td>Retail Operations</td>
<td>30507 Irvington</td>
<td>Brown, Paul</td>
<td>323-33-2345 A</td>
<td></td>
</tr>
<tr>
<td>Retail Operations</td>
<td>32544 Massapequa</td>
<td>Hamilton, Gavin</td>
<td>324-54-3456 A</td>
<td></td>
</tr>
</tbody>
</table>

Other Tests:

- Test for no Withholdings
- No Vacation or Sick Days
- Invalid address
Testing Examples: Duplicate Vendors and Payments

- **Entertainment Company**
  - 17% duplicate vendors
  - $500,000 in recovered duplicate spend
  - $750,000 in duplicate spend

- **Computer Manufacturer**
  - 5% duplicate vendors
  - $16M in potential duplicate spend

- **Information Services Company**
  - 9% duplicate vendors
  - $1.6M in possible duplicate spend
### Testing Examples: Duplicate Payments

<table>
<thead>
<tr>
<th>VENDOR_NAME_ NORMALIZED</th>
<th>VENDOR_NAME</th>
<th>VENDOR_NBR</th>
<th>INVOICE_NBR</th>
<th>CHECK_NBR</th>
<th>INVOICE_DATE</th>
<th>DISTRIB_AMOUNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHARMERICA</td>
<td>+PHARMERICA</td>
<td>0000000356</td>
<td>411316711119</td>
<td>23345059</td>
<td>11/19/2001</td>
<td>5,323.16</td>
</tr>
<tr>
<td>PHARMERICA</td>
<td>PHARMERICA</td>
<td>000664457</td>
<td>411316711119</td>
<td>23375682</td>
<td>11/19/2001</td>
<td>5,323.16</td>
</tr>
<tr>
<td>WESTERNFOODSINC</td>
<td>+WESTERN FOODS INC</td>
<td>000000052</td>
<td>378158</td>
<td>23247120</td>
<td>9/6/2001</td>
<td>3,828.95</td>
</tr>
<tr>
<td>WESTERNFOODSINC</td>
<td>WESTERN FOODS INC</td>
<td>000912302</td>
<td>378158</td>
<td>23266700</td>
<td>9/6/2001</td>
<td>3,828.95</td>
</tr>
<tr>
<td>WESTERNFOODSINC</td>
<td>+WESTERN FOODS INC</td>
<td>000000052</td>
<td>363172</td>
<td>23119489</td>
<td>6/26/2001</td>
<td>3,394.65</td>
</tr>
<tr>
<td>WESTERNFOODSINC</td>
<td>WESTERN FOODS INC</td>
<td>000912302</td>
<td>363172</td>
<td>23129197</td>
<td>6/26/2001</td>
<td>3,394.65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VENDOR_NAME_ NORMALIZED</th>
<th>VENDOR_NAME</th>
<th>VENDOR_NBR</th>
<th>INVOICE_NBR</th>
<th>CHECK_NBR</th>
<th>INVOICE_DATE</th>
<th>DISTRIB_AMOUNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HALO</td>
<td>HA LO</td>
<td>000721207</td>
<td>827566</td>
<td>23336819</td>
<td>9/6/2001</td>
<td>947.55</td>
</tr>
<tr>
<td>HALO</td>
<td>HA LO</td>
<td>000591572</td>
<td>827566</td>
<td>23298033</td>
<td>9/6/2001</td>
<td>947.55</td>
</tr>
<tr>
<td>GOLDBERGANDSOLOVY</td>
<td>+GOLDBERG AND SOLOVY</td>
<td>000000054</td>
<td>826033</td>
<td>23412062</td>
<td>2/12/2002</td>
<td>942.32</td>
</tr>
<tr>
<td>GOLDBERGANDSOLOVY</td>
<td>GOLDBERG AND SOLOVY</td>
<td>000278066</td>
<td>826033</td>
<td>00000000</td>
<td>2/12/2002</td>
<td>942.32</td>
</tr>
</tbody>
</table>
Payroll Fraud

- Databases involved
  - Payroll
  - Personnel
  - Time & attendance
  - Employee directory
  - Terminated employee directory
  - Valid SSN listings
  - Leave and vacation
  - Employee expenses
Possible Tests for Payroll Fraud

- Duplicates:
  - Payees on same date
  - Same or similar names
  - SSNs
  - Addresses
  - Telephone numbers
  - Direct deposit account numbers
  - Work locations
  - Work telephone numbers
Possible Tests for Payroll Fraud

- Comparisons: Paycheck, but ...
  - no time & attendance
  - not in employee directory
  - no leave ever taken
  - no expenses
  - terminated employee
  - no pension or other deductions
  - no valid SSN
  - no, or scarce, personnel data
Possible Tests for Payroll Fraud

- Employees with P.O. Box addresses
- Wages inconsistent with job classification
- Payroll disbursements above certain amounts (by pay period or cumulative)
- Premium pay above certain amounts
- Benefit contributions but not on payroll
- Pay date precedes employment date
- Overtime with other activity indicators
Purchasing Fraud

- Databases involved
  - Vendor
  - Personnel
  - Dun & Bradstreet
  - Employee directory
  - Terminated employee directory
  - Employee expenses
  - Inventory
  - Accounts payable; accounts receivable
Possible Tests for Purchasing Fraud

- Duplicate disbursement amounts
- Duplicate invoice numbers/dates
- Duplicate disbursements on same date
- Disbursement to vendor not in vendor database
- Vendor name/address/phone # same as employee name/address/phone #
- Vendor name *similar* to employee name
Possible Tests for Purchasing Fraud

- Purchases inconsistent with inventory
- Missing purchase request (PR) #, purchase order (PO) #, receiving report (RR) #
- Names on PR, PO, or RR missing; or the same; or not in employee database
- Vendor address is P.O. Box
- Invoice’s “pay to” address different from address in vendor or contract database
Possible Tests for Purchasing Fraud

- Vendor with no telephone number or contact name
- Invoice numbers from same vendor in unbroken sequence
- Invoice amounts just below higher-level-approval-required threshold
- Multiple vendors with same address, telephone, contact person
- Vendors with similar or similar-sounding names
Possible Tests for Purchasing Fraud

- Unit prices rising rapidly or inconsistent with historical prices
- Unit prices for same items inconsistent among different vendors
- Matches of item numbers/description being purchased and sold for scrap
- Inventory level fluctuations inconsistent with production or sales
- Contract or PO “ship to” address matches employee address
- Contract or PO “ship to” address does not match any company site addresses
- ?
- ?
Data Analysis – A Generic Approach

- Identify all available databases
  - Internal to the organization
  - External to the organization
- List record fields in all available databases
- Formulate hypotheses about record field relationships
- Program analytical tests for each hypothesis
- Run tests (output is your “hit list”)
- Evaluate initial hit list and refine the tests
- Re-run refined test to produce shorter, more meaningful hit list (repeat steps 5-7, as needed)
Data Analysis – A Generic Approach

- Evaluate (via record analysis, interview, or other technique) every item on the refined hit list
- Dispose of every hit:
  - Valid explanation found
  - Probable improper transaction – full investigation needed
- Identify control problems and corrective actions needed
Data Analysis

- Keep the programs developed and refined over time
- Run as batch / script programs as frequently as resources allow
- Explore feasibility of *real-time* tests of transactions
Real-Time Transaction Tests

- Add a loop to the existing accounting system transaction process
- Loop goes through the internal or external audit organization
- Transactions that fail the audit organization’s [secret] tests get stopped for further analysis
Identify data patterns indicative of fraud.
- Develop Fraud Criteria
- Scan data at short intervals.
- Trigger alarms when data meets criteria.
- Data tests can be continuously refined.

CAATs in Continuous Monitoring

- Identify Tests to detect Anomalies
- Continue to refine tests
- Short interval data feeds from IT into a secure audit folder
- If anomalies detected, yes, e-mail exception reports
- No, continue to refine tests
Benford’s Law

- A Digital Analysis Technique
- The first digits of numbers are not randomly distributed
- Distribution of first four digits:
  - 1 – 30.1%
  - 2 – 17.6%
  - 3 – 12.5%
  - 4 – 9.7%
Benford’s Law

![Graph showing the distribution of first digits according to Benford's Law]
Benford’s Law

- So what?
- Use Benford’s Law to identify anomalous transactions in:
  - Accounts payable and expenses
  - Accounts receivable and sales
  - Refunds
  - Payroll
  - Estimations in the general ledger
Purchase Order Amounts

First Digit

Frequency

1 2 3 4 5 6 7 8 9
Bank Deposit Amounts

First Digit

Frequency
Benford’s Law

- Invented, contrived, or manipulated numbers do not follow Benford’s Law
- Doesn’t apply to all data sets
  - e.g. sets of data with numbers that have imposed ceilings or floors such as IRA contributions or hourly wage rates
- Once again, however, human analysis of anomalies is needed
Financial Crime Investigator

- Artificial intelligence or expert system
- Relational database containing:
  - Fraud indicators
  - Fraud schemes
  - Fraud statutes and elements of proof
  - Investigative steps
  - Steps for database queries to detect specific schemes
If you have a fraud indicator, FCI will list possible fraud schemes for that indicator.

If you suspect a fraud scheme, FCI will:
- List other indicators to look for
- List fraud statutes under which the scheme can be prosecuted

If you identify a potential fraud statute violation, FCI will list the elements of proof needed to prove the crime.
Financial Crime Investigator also has:

- Decision trees for opening, conducting, continuing, and terminating an investigation
- Checklists for making decisions at any point in your case development
- Library aids: interviewing techniques, using the case theory approach, case tactics and strategy
- Guidance for writing investigation reports
- Glossary of terms (hypertext)
- "How-to" guides (e.g. proof of cash; file a complaint/get a case prosecuted; prove intent; find mail-drop addresses; etc.)
- Key sources for additional data such as public records; credit information; bankruptcy/judgment/lien information; vehicle registrations; business background information; etc.
Case Study: Barings Bank

- Nick Leeson was a 27-year-old rogue derivatives trader in Barings’ Singapore office.
- He single-handedly collapsed the 233-year-old bank by trading with bank funds.
- He reported his gains (and got large bonuses as a result); but hid his losses in a “5-eights” account (account number 88888).
- The bank collapsed in Feb. 1996.
- In Dec. 1995, account 88888 had $512 million in losses hidden in it.
- Among the means Nick used to conceal the losses in 88888 was “suppression” of the account in printed reports to London.
- Could data analysis have saved Barings?
Case Study: Barings Bank

- Simply totaling – electronically – the debits and credits in the reports to London would have revealed the “suppression”
- Account 88888 was supposed to be a low-balance adjustment account for minor trading errors; hence, a simple stratification test would have revealed the anomaly
Case Study: The Fool Service Motor Pool

- Whatsamatta University operates a motor pool to service and maintain its large fleet of vehicles (cars, trucks, and research vessels)
- Arthur Fulldodger has managed the motor pool for 25 years, and everyone agrees he does a terrific job
- Anne Alasiss, CPA, is assigned to test motor pool controls and expenses as part of the annual audit
- Anne notes that the manager, Art, seems to be involved in every aspect of running the motor pool and keeping its service records
- Anne downloads the service pool’s database of service records and conducts a number of tests using IDEA
- Anne uses vehicle mileage records and gasoline charge card records to calculate vehicle miles per gallon
Case Study: The Fool Service Motor Pool

- Anne finds that most vehicles appear to be getting very poor gas mileage—less than 10 miles per gallon.
- A more detailed, monthly vehicle-by-vehicle analysis reveals that some vehicles get very good mileage (30+MPG) in some periods and very poor mileage (<5 MPG) in others.
- Further investigation (including surveillance) reveals that faculty and students fill their vehicles at the motor pool.
- Art charges them $.50/gal; pockets the money; and records the gas charge against university vehicles.
- Anne also extracts and plots repair costs by vehicle on a month-by-month basis.
- This analysis shows several anomalous results.
Case Study: The Fool Service Motor Pool

- New vehicles seem to have an unusual number of repairs—even vehicles still under warranty
- Graphs show a peak of repair activity in the two months prior to the end of each vehicle’s service history
- What did Anne’s further investigation reveal?
- Art is repairing cars for faculty and students at “deep discount” rates
- Art is refurbishing fully-depreciated, “no-more-useful-life-left” vehicles just before he “retires” them by selling them to his pal, Rudolph Mentari, who happens to own “Honest Rudy’s” Used Car Dealership
rabbids eat grass... grass grows in soil... rain falls from clouds... wolves eat rabbits... etc... etc...

if I kill off all the wolves, the rabbits will eat up the grass, & the soil will all wash away!

INFORMATION...

KNOWLEDGE...

...WISDOM!

Source: http://www-personal.si.umich.edu/~nsharma/dikw_origin.htm
And now ...
Using Computer-Assisted Auditing Techniques to Detect Fraud
Appendix

Data Analysis Techniques: Detailed Commands
Filters

- Show me only certain transactions that I might be interested in
- I.e. filter out the “static” inherent in large databases
- E.g. from the database of all invoices paid this month, show me those with invoice dates more than 30 (or 60, or 90) days prior to this month
Sorting

- Put the data into an order that is easier to analyze
- E.g. aged accounts payable
- Sort by vendor name;
- Or sort by vendor address;
- Or sort by vendor telephone number;
- Or EFT account number
Statistics

- Average value, standard deviation, highest/lowest value, etc.
- Show me all vendor payments that are more than two standard deviations from the mean
- Show me every unit price for product A that is more than 1 standard deviation from the mean
Gaps

- Used when we have pre-numbered forms or transactions
- Show me all missing checks
- Show me missing health insurance claims
- Show me missing purchase orders
- Show me missing vouchers for benefits
Duplicates

- Used in situations when duplicates should not occur
- Show me all payroll transactions in the pay period that have duplicate:
  - Payees
  - SSN
  - Payment addresses
  - EFT transfer destinations
We all use an aged A/P listing

AGE function can calculate the number of days between two date fields

Show me the time elapsed from when an item was purchased and placed in inventory and when it was disposed of as “scrap” or “due to obsolescence”

Show me the time elapsed between award of a major contract and the contracting officer’s purchase of a condo in the Bahamas
Expressions and Calculations

- Can be used to test for and detect errors in accounting software
- But, can also be used to detect outside intervention into an otherwise okay system
- Recalculate units x unit price and show me all cases where the result does not equal the value in the extended amount field
- Recalculate vacation accruals for every payroll transaction and show me every case where the result differs from what was recorded
Classify

- Counts the number of unique values in a selected character field(s) and the corresponding totals of other numeric fields
- Show me how many hysterectomy procedures have been performed, by sex:
  - Female: 127
  - Male: 3
  - [sex field blank]: 12
Summarize

- Similar to CLASSIFY, but sorts data by specified field and provides a detailed listing of transaction information for all records in that classification
- Show me all transactions for vendor A, B, C
- Show me all payroll transactions for employee A, B, C
- Show me all employee expenses transactions for employee A, B, C
Stratify

- Groups transactions into specified ranges of values (strata)
- Show me numbers of contracts within certain strata
  - $0 - $45,000: 27
  - $45,001 - $49,000: 12
  - $49,001 - $50,000: 425
  - $50,001 - $100,000: 5
  - $100,001 - $1,000,000: 3
Join and Define Relation

- JOIN combines data from two databases into a single database
- DEFINE lets you specify the relationship of interest
- Show me employee expense claims by employees who were on vacation/sick/holiday leave on the day the expense was incurred
- Show me all instances in which a person made a DNC/RNC campaign contribution within 60 days of getting a Federally-guaranteed loan
Trend/Regression Analysis

- Trend analysis looks at historical trends that data exhibit
- Regression analysis uses historical data to predict what future values will be
- Looking at electricity usage trends will show anomalous changes
- Regression analysis of electricity usage will predict what the usage should be in the current period
Appendix

Where to Get More Information
Where to Get More Information

- About ACL: www.acl.com
- About IDEA: www.audimation.com
- About Detective Toolkit, Fraud Investigator, and Similarity Search Engine: www.infoglide.com
Where to Get More Information

- About ViCLAS: www.mtps.on.ca/Year/ViCLAS
- About Data Mining:
  - www.gartner6.gartnerweb.com
  - www.statserv.com/datamining.html
  - www.datamining.org/sites.htm
  - www.wizsoft.com
- About Financial Crime Investigator:
  www.cci2.com/fci_prod.htm
Vinesh R Bulsara
vinesh.r.bulsara@us.pwc.com

Vinesh is a Director in PricewaterhouseCoopers’ Process Improvement practice specializing in Data Management (DMG). He has over nine (9) years experience in performing IT reviews (security, data mining, project management), and process and controls review. His experience including auditing Entertainment and Media, and Consumer Products, and Technology companies. During the past two years, Vinesh has focused on Data Management in support of internal and external clients with SAS 99 and other automated testing.