

Using Data Analytics to Detect Fraud

Gerard M. Zack, CFE, CPA, CIA, CCEP
President, Zack, P.C.

Introduction to Data Analytics

Course Objectives

- How data analytics can be used to detect fraud
- Different tools to perform data analytics
- How to walk through the full data analytics process
- Red flags of fraud that appear in the data
- Data analytics tests that can be used to detect fraud
- How to analyze non-numeric data, such as text and timelines, for signs of fraud

Introduction to Data Analytics

- *Data analytics*, as it applies to fraud examination, refers to the use of analytics software to identify trends, patterns, anomalies, and exceptions within data.

Introduction to Data Analytics



- Especially useful when fraud is hidden in large data volumes and manual checks are insufficient
- Can be used reactively or proactively

Introduction to Data Analytics

- Effective data analysis requires:
 - Translating knowledge of organization and common fraud indicators into analytics tests
 - Effectively using technological tools
 - Resolving errors in data output due to incorrect logic or scripts
 - Applying fraud investigation skills to the data analysis results in order to detect potential instances of fraud

Introduction to Data Analytics

- Data analysis techniques alone are unlikely to detect fraud; human judgment is needed to decipher results.



Types of Data That Can Be Analyzed

- Relevant data comes from numerous sources and takes numerous forms
 - Accounting and financial data
 - Human resources data
 - Customer data
 - Vendor data
 - Internal communications and documents
 - External benchmarking data

Types of Data That Can Be Analyzed

Structured data

- Sales records
- Payment or expense details
- Payroll details
- Inventory records
- Financial reports
- Found in accounting software, databases, spreadsheets, etc.

Unstructured data

- Email and instant messages
- Payment text descriptions
- Social media activity
- Corporate document repositories
- News feeds

Types of Data That Can Be Analyzed

- Big data:
 - **Gartner:** “High volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision-making, insight discovery, and process optimization.”
 - Information of extreme size, diversity, and complexity that will eventually tell a story once all the data sets and pieces from different sources are connected.

Population Analytics

- Although testing a sample of data is a valid *audit* approach, it is not as effective for *fraud detection* purposes.
- To detect fraud, data analysis techniques must be performed on the full data population.
- Need to define population boundaries, including amount of historical data to include.

Benefits of Data Analytics

- Increase efficiency and effectiveness.
- Boost productivity and profitability.
- Reduce sampling errors.
- Assess and improve internal controls.
- Revise or reinforce policies.
- Monitor trends.

Benefits of Data Analytics

- Identify fraud before it becomes material.
- Focus detection efforts on suspicious transactions.
- Gain insight into how well internal controls are operating.
- Compare data from diverse sources to identify instances of fraud or noncompliance.

Challenges in Using Data Analytics

- Poorly defined scope
- Data acquisition
 - Manually maintained data
- False positives
- Lack of familiarity
 - Data storage systems
 - Software systems
 - Organizational processes

Challenges in Using Data Analytics

- Data security and integrity concerns
- Privacy and confidentiality concerns
- Evolving business processes/activities
- Learning curve
- Cost of data analytics software
- Evolution of fraud schemes
- Element of concealment

Data Analytics Software

- Commercial data analysis programs
- Spreadsheets/databases
- Customized data analysis programs
- Shareware/freeware data analysis programs

Spectrum of Analysis

- Ad-hoc testing
- Repetitive testing
- Continuous testing

