Understanding Your Data and Best Practices in Data Analytics

Lance E. H. Schmidt, CPA, CFE, Principal

Learning Objectives

This course will feature a live demonstration of how to understand your data utilizing the program IDEA. In addition, this course will cover best practices in performing data analytics using IDEA. Specific learning points include:

- Understanding your data to decrease time and increase effectiveness
- Normalizing your data for efficient analysis
- Basic data analytic methodology
- Importance of completeness in data analysis
- Best practices in data analytics

CBOK 2015 Practitioner Survey

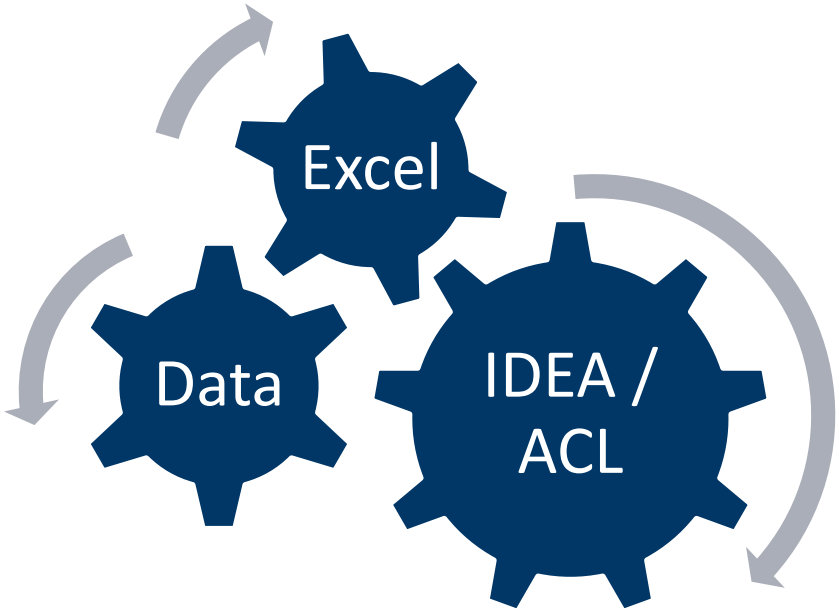
- Only 40% of CAEs say their use of technology is “adequate”
- 20% of CAEs say their departments rely primarily on manual testing
- 53% of audit departments use a tool for data mining or data analysis
- 80% of CEOs (Management) say data mining and analysis is strategically important to their organizations

Source: www.theiia.org/goto/CBOK

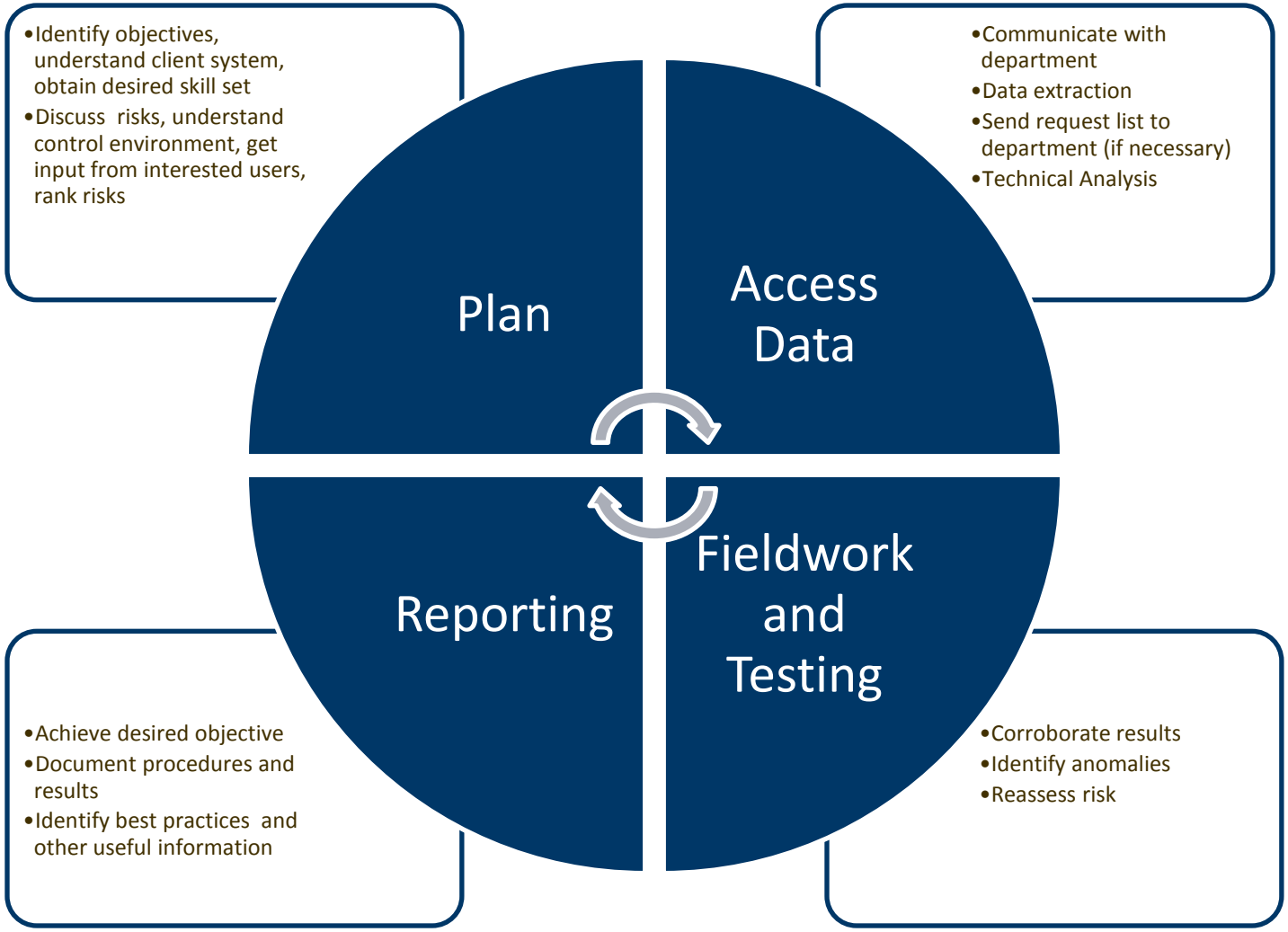
Quick Poll

- How well is your team utilizing *Computer Assisted Audit Techniques* (CAAT)?
- What tools are being used for data analysis?
- What is keeping us from performing *better* data analysis?

It is difficult to explain what data analysis is, but you know it when you see it.



Basic Data Analysis Methodology



Access Data

- Do not just settle for “No”
  - The data exists, we need to find a way to obtain it
- Consider IT personnel involvement
  - The accounting personnel may not be the best resource
  - Save time by getting data in the correct format
- Perform a cost benefit analysis
- Excel is not always the best format

Access Data (continued)

- Commonly requested information types include, but are not limited to:
  - General Ledger Detail
  - Journal Entries
  - Accounts Receivable
  - Accounts Payable
  - Contribution Records
  - Cash Disbursements
  - Procurement Records
  - Payroll
  - Credit Cards
  - Travel and Entertainment Expenses
  - Subsidiary Ledgers (Various)
  - Along with the above noted data sets, the following information is also needed to adequately interpret the data files:
    - ◊ Chart of Accounts
    - ◊ Data Dictionary
    - ◊ File Source Listing – indicating from which systems and locations the data was extracted
  - Vendor Master File (A/P, A/R, Sales)

Technical Data Analysis

Ensure proper import (anyone)

Test data for completeness (anyone)

Normalize data and prepare for analysis (understanding of “data” and types of tests to be performed)

Design analytics to address identified risk and accomplish objectives (team effort)

Proper Import

- The 2nd most important step in the data analytic process
- Document source of data (screen shots)
- Is your source data format ideal?
- Utilize field statistics
- Ensure data field types are correct
  - 4 data types in IDEA

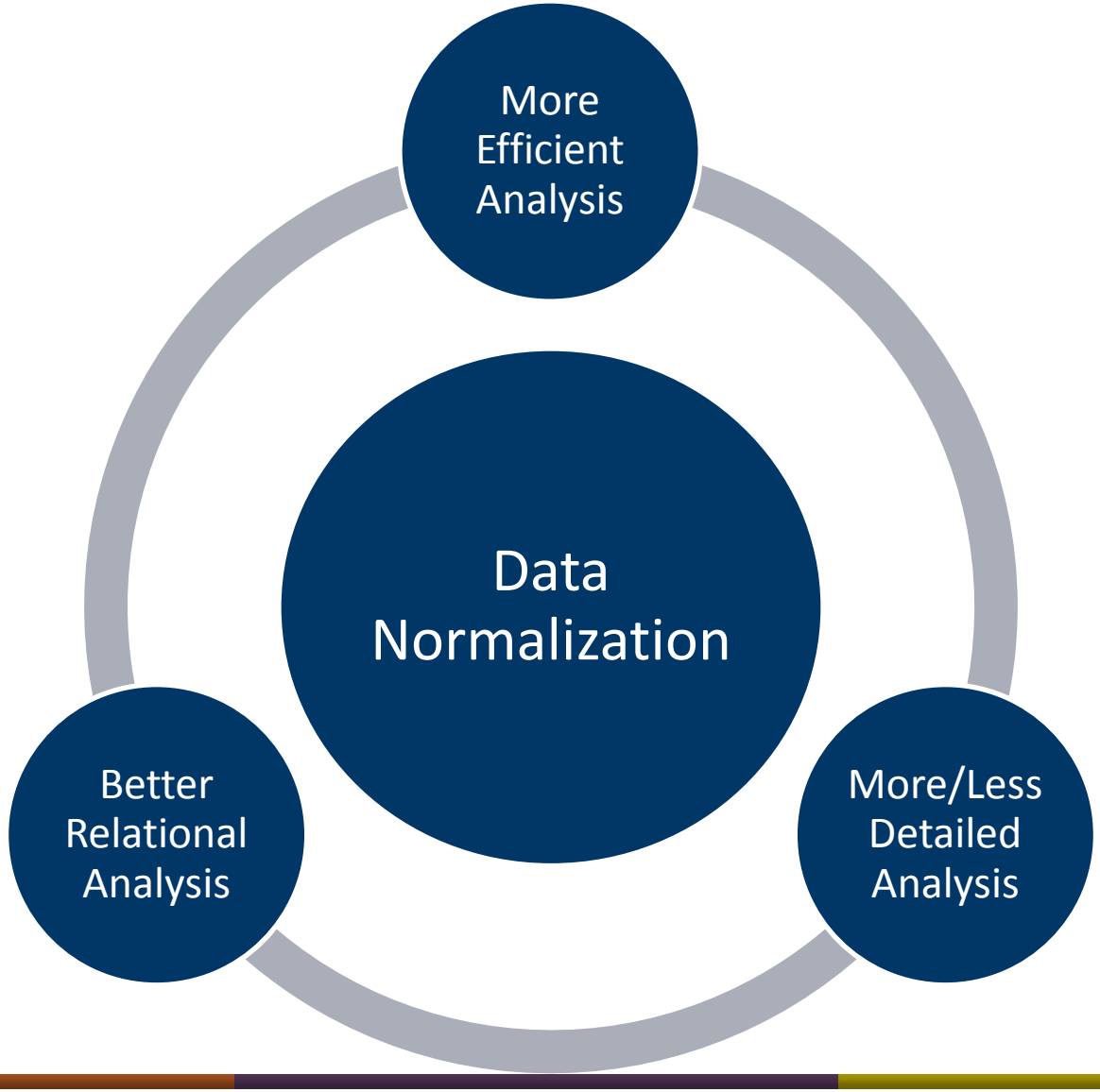
4 Types of Data in IDEA

- IDEA data types
  - **Character** – text data; this should be used for included data that is not, Numeric, Date, or Time. Also, if you are unsure whether a data field maintains the same structure throughout the entire database, character should be used as the "default" selection.
  - **Numeric** – numeric data; generally, the rule of thumb is that if you are going to use (or may use) the field as a part of a mathematical calculation, it should be defined as Numeric, the field should contain only numbers (and commas, parenthesis, periods, in some cases IDEA will recognize "DR" and "CR").
  - **Date** – date data (a mask is required, such as "MM/DD/YYYY"); if the data field contains date information with a consistent mask throughout, it should be imported as a date.
  - **Time** – time data (a mask is required, such as "HH:MM:SS"); if the data field contains time information with a consistent mask throughout, it should be imported as a time.

Completeness

- The most important step in the data analytic process
- Financial data
  - Agree to general ledger or trial balance
- Operational data
  - Report totals

Data Normalization



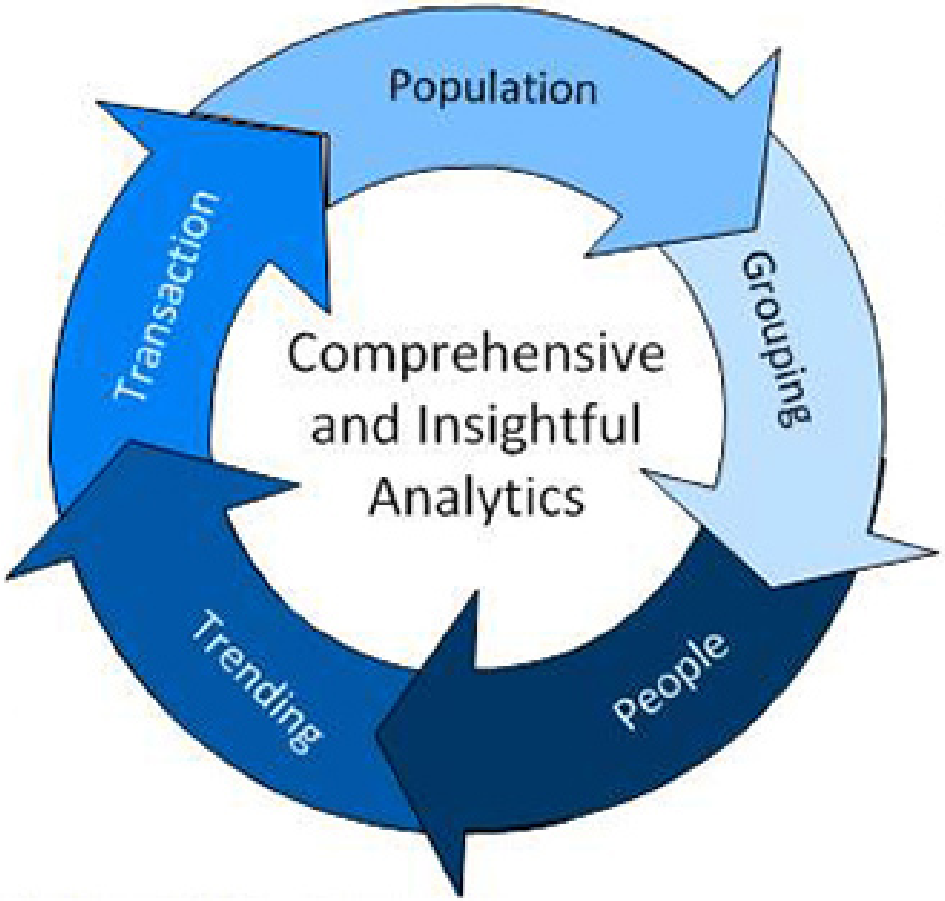
Common Data Normalization Fields

- Absolute Value
- Month
- Year
- Month Year/ Year Month
- Account Number Segmented
- Fund (parent and child fund)
- Fund Type and Fund Subtype
- Financial Statement Class
- Subsidiary Ledger Identifier
- Functional expense allocation
- Workday / Weekend
- Day of the week
- Days to posted
- Balance Sheet/Income Statement
- Single Audit/ UGG Compliance
  - CFDA / CFSA Number and Description
- Shortened version of account number or vendor field
- Statistics by vendor/line item description

Data Normalization Now =

Significant Time Savings Later

Technical Data Analysis – Design Analytics



Types of Analytics

- Population analytics — Understand the entire population.
  - Example test: Stratification by $10,000 increments
- Grouping analytics — Summarize transactions into meaningful groups.
  - Example test: Summarization of manual journal entries by accounts
- People analytics — Garner insight into who benefits from and who is responsible for a transaction.
  - Example test: Count and sum of journal entries by inputting user

Types of Analytics (continued)

- Trending Analytics — Show results through time.
  - Example test: Sum cash disbursements by month
- Transaction Analytics — Isolate transactions exhibiting particular traits.
  - Example test: Transactions occurring on a weekend

Hands on Example

Understanding your data

Interpret the Results

- Team effort
- Don’t lose sight of the overall objective
  - Don’t get “lost in the data”
- What value added items come to light when looking at the results?
- What new questions can we ask?
- How can we better ask analytical questions?

Top 10 Underutilized Features of IDEA

Top 10 Underutilized Features of IDEA

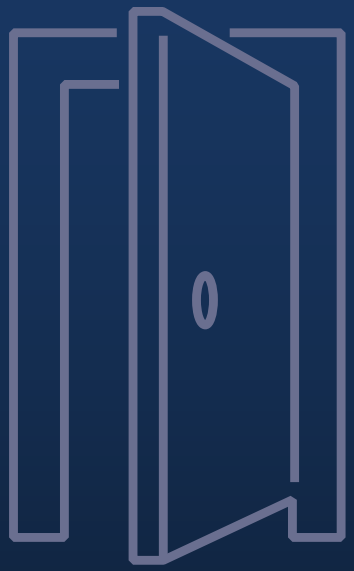
1. Summarization (similar to subtotal and pivot table)
2. Re-run
3. Right click, display all (similar to filtering)
4. Search (wildcard and proximity)
5. Join (similar to vlookup)

Top 10 Underutilized Features of IDEA

6. Append
7. Define action field
8. Stratification
9. Editable text fields
10. Boolean and Multistate fields

Recap: Implementation Considerations

- Training
- Director/manager involvement
- Plan to use of data analytics
- Perform data analytics early on
- Budget considerations
- Do not settle for data that is not user friendly
- Save import definitions, scripts, and commonly used formulas
- Share best practices, tips, and tricks



Questions?

Lance Schmidt, CPA, CFE, Principal
State and Local Government
Direct 863-680-5634
lance.schmidt@CLAconnect.com

Let us know how we can help!

CLAconnect.com