Data Analytics for State Auditors

A Joint Presentation by the North Carolina Office of the State Auditor and RSM

41st annual conference of the National State Auditors Association

Cherokee, NC - June 12, 2019
With You Today

Dan Boulden
• NC Office of the State Auditor
• Supervisor, Data Analytics
• Email: dan_boulden@ncauditor.net

Jim Tarantino
• RSM, National Data Analytics Practice
• Manager, Data Analytics & Automation
• Email: jim.tarantino@rsmus.com
Key Topics / Agenda

- Data Analytics (DA) Defined
- Drivers for Use of DA
- Setting Up a DA Function
  - Personnel
  - Projects
  - Data
  - DA Model
- Q&A
What are Data Analytics (DA)?

Using technology to automate, in part or in whole, the discovery, interpretation and communication actionable insights and meaningful conclusions derived from data.

**Digital Evidence** + **Algorithms** + **Technology**
- Digital Evidence: (business data)
- Algorithms: (set of operations)
- Technology: (understood by computers)

<table>
<thead>
<tr>
<th><strong>Automated</strong></th>
<th><strong>Purpose</strong></th>
<th><strong>Source of Rules</strong></th>
<th><strong>Focus</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ad hoc</td>
<td>Descriptive</td>
<td>Experts</td>
<td>Risk (KRI)</td>
</tr>
<tr>
<td>Repeatable</td>
<td>Diagnostic</td>
<td>Statistics</td>
<td>Performance (KPIs)</td>
</tr>
<tr>
<td>Continuous</td>
<td>Predictive</td>
<td>Machine Learning</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- Automated
  - Ad hoc
  - Repeatable
  - Continuous
- Purpose
  - Descriptive
  - Diagnostic
  - Predictive
  - Prescriptive
- Source of Rules
  - Experts
  - Statistics
  - Machine Learning
DA is a Part of Digitally Transforming Audit

Answer questions about past, present, and future
- Bulk calcs, summarization and profiling
- SoD, business rules, and expert rules
- Data visualization
- Risk scoring, modeling, and statistics
- Process mining, text mining
- Unstructured data processing
- Machine Learning, Deep Learning and AI

Automate and routinize key audit tasks
- Report generation & distribution
- Low cognitive tasks
- Manual, repetitive or high volume tasks
- Scheduled jobs and queries
- Cross-application “macros”
- Higher-order task automation (with AI)

Organize, prioritize and deliver on audits
- Quick sprints, adaptable to changes
- Incremental work vs all at once
- Increase information & communication flow
- Client and peer collaboration
- Risk backlog vs defined plan
- Alerts, notifications and workflows
Why the Renewed Interest in DA?

- Unprecedented velocity of change and risk
  - Service offering and process changes
  - Technology transformation
  - Workforce changes
  - Geographic coverage / expansions
  - Regulatory landscape

- Data is critical to business intelligence
- Data volume, variety, velocity, and veracity
- DA becoming standard practice for regulators
- Expectation that audit to provide more value
- Opportunity for audit to challenge the status quo

Increased pressure to focus on more sophisticated and relevant risks while still responsible for traditional risks and compliance areas.

“Datafied” processes and commoditized automation and analytical capabilities create a RENEWED opportunity to scale audit and compliance efforts.

Audit must scale and adapt
What does DA mean for the State Auditor?

**Efficiency & Effectiveness**
- Expedited activities and resources
- Avoid unnecessary work

**Coverage & Accuracy**
- Focus on the “right” risks and controls
- Enable more substantive and comprehensive reviews

**Insight**
- Understand of business operations
- Aware of relevant changes
- Substantiate impact of audit findings and remediation efforts

**RESULTING IN**
- Robust business rules and logic
- Integrated data from multiple sources
- Examination and profiling of full populations
- Data informed planning and risk assessments
- Prioritized focus based on risk scoring and risk-based sampling
- Data visualization and dynamic reporting
- Continuous monitoring of persistent risks

DA will aide in your pursuit to protect the interests of the taxpayers of your state.
DA Can Be Implemented in a Variety of Models

Auditor Self-Service
- Ad hoc analysis

Data Retrieval / CAATs Team
- Data Delivery
- Test Automation
- Audit Support

Data Analytics (DA) Team
- Risk ID
- Triage
- Audit Generation

Our Focus
Today
Setting Up a Specialized DA Division in Your Office

The tools you will need...

• Personnel
• Projects
• Data
• DA Model
Setting up a DA Division: Personnel

• How to avoid needing a “unicorn”
• Getting the most value from your audit staff
• An Auditor or a Programmer
  • Audit analytics engineer
    • Auditor position
    • Rotation based on project
  • Analytics programming experts
    • Non-audit position
Personnel: How to Avoid Needing a “Unicorn”

**Challenge:** DA requires skills, knowledge and experience from multiple domains

Expert auditors who are also expert programmers, and an expert on each agency, would be wonderful, but not practical in any way

**Solution:** Split the positions up so that you can have one or more expert in each area

- This allows you to get the most value in each area
- An auditor does not need to be an expert programmer and visa versa
Personnel

Auditors do not need to be expert programmers (and visa versa)

• Your auditors work with management to define the required analytics

• Then, the auditors can map out (storyboard) the DA and work with the programmer to achieve the DA

• This way the programmer does not need to fully understand auditing as they have an auditor as a guide to keep them on the correct path for the DA
Personnel: An Auditor or a Programmer?

• Auditing Analytics Engineers

• Analytics Programming Experts
Personnel: An Auditor or a Programmer?

Auditing Analytics Engineers (Auditor)

• Your auditors who work with management to understand the DA requirements
• Storyboard the DA so that both management and the programmers can understand
• Work with the auditee to endure the data is valid
• Work directly with the Analytics Programming Experts to ensure that the DA is programed correctly
• Test the DA output
• Have a base understanding of the coding language*

*for permanent DA Auditing Analytics Engineers
Personnel: An Auditor or a Programmer?

**Analytics Programming Experts** (non auditor position)

- Be an expert in the coding language
- Write the DA
- Work closely with the Auditing Analytics Engineers to ensure they are writing the code to correctly accomplish the DA
Personnel: Mesh Auditors, Programmers (and DBA’s)

Audit Client

Audit Team
- Audit Liason
  - Consult and Assist
  - Provide Guidance
  - Process, System Knowledge
  - Connections with Key Contacts / SMEs

Initiate Performance Audit or Investigation

Ideas / Opportunities

Audit Liason

DA Team
- Auditing Analytics Engineer
  - Understands complex Analytics
  - Translates audit requirements into technical requirements
  - High-level design via Story Board

- Analytics Programming Experts
  - Develop Recurring Analytics / Utilities / Dashboards

- Database Administrator (as needed)
  - Data engineer / Wrangler

Analytics Programming Experts

Expertise

Storyboards / Results

Analytic Tools

Standardized Data
Rotation of Auditors

- Gaining an understanding of the
  - The agency
  - The agency's staff
  - The agency's data

- Very time consuming

- Rotate in or heavily rely on audit staff that are already fully trained on the agency
  - Speeds up DA
  - Avoids annoying the agency
Selecting DA Projects

- Sophistication vs sustainability
- Time to create vs value
- The power of a clear scope
- The problems with an unclear scope
Selecting DA Projects: Key Considerations

- **Sophistication vs Sustainability**
  - X Target
  - Do not create

- **Value vs Time to Create**
  - X Target
  - Do not create
Selecting DA Projects: Major Roadblocks

- Vague DA project definitions
- Data Quality
- Delays from Auditee
- The DA team is pulled in different directions
Selecting DA Projects: The agony of an unclear scope

Vague or grand ideas will:

• Waste time and effort

• Frustrate everyone

• Return unintended results
Selecting DA Projects: The Agony of Unclear Scope

- Inefficient and out of focus DA’s
- Scope creep
- Fishing / confirmation bias
- DA group goes in their own direction
- Status meetings are chaotic and frustrating
- End results are a surprise
  - The DA is looking at the wrong things
Selecting DA Projects: Unclear Scope Examples

- “Look at this data and find any issues”
- “I feel there is something wrong at this agency, see what you can find”
- “Look for issues in the data”
- “See if you can find data that shows if there is an issue”
Selecting DA Projects: Power of a Clear Scope

• When selecting DA projects

  • Select projects that are able to have
    • Clear and obtainable scope and objective(s)
    • Feasible analytical approach and procedure(s)
    • Available, complete and accurate data
    • A new or better insight than status quo

“Every successful data science project begins by clearly defining the problem that the project will help solve.”

- John D. Kelleher and Brendan Tierney “Data Science”
Selecting DA Projects: Scoping Questions

Events can be **static** one time occurrences, **repeated** conditions, or **dynamic** changes (e.g. flux)

<table>
<thead>
<tr>
<th>Who</th>
<th>What</th>
<th>Where</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Actors involved in the event</td>
<td>• A label, flag, or ID of the event</td>
<td>• Location of the event</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>When</th>
<th>Why</th>
<th>How</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Time of the event</td>
<td>• Trigger or cause for the event</td>
<td>• Means by which the event took place</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How Much</th>
<th>How Often</th>
<th>Which</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Magnitude of the event</td>
<td>• Frequency of the event</td>
<td>• Features to facilitate choice or selection</td>
</tr>
</tbody>
</table>

Example: “Identify food stamp vendors (**who**) in any location receiving NC benefit dollars (**where**) that over the last 5 fiscal years (**when**) cashed food stamp dollars (**what**) to a pool recipients frequently traveling unusual distances for their benefits (**which**).”
Selecting DA Projects: Collaboration Is Key

Objective
“Voice of the Client”

Risk and Control Challenges Within Auditees

Detect
Deliver

Auditors

DA Specialists

Approach
Collaboration Zone

Syntax

Result

Interpretation

Procedure
Selecting DA Projects: Power of a Clear Scope

• Have Regular status meetings

  • Ask if there are any roadblocks that you can help get past

  • Go over the scope and insure that the team is on target
    • Sometime the data shows valid reasons to alter the scope
      • Be open to changes in the scope

• Understand that most valuable DA projects will run into roadblocks and dead ends
  • If it was easy, someone would have already done it
  • Adjust timelines as needed

• Do not rush a DA just to meet a deadline
Data Quality

Getting good data may be the hardest part of any DA

what a mess!
Delays from Auditee

• Auditee Does not know how to get the data to you

• Auditee is actively slow

• Auditee lawyers up
Delays from Auditee

- Auditee Does not know how to get the data to you
  - Quickly find out if there is someone better to talk to about the data
    - Locate and understanding the data (Auditor and/or key data user)
    - Capable of handling the volume and variety of data (DBA)
  - Often the data should come from the auditee’s database administrator and not the data user
Delays from Auditee

• Auditee is actively slow

  • Have clear timeline and path for escalation
  • Data is critical - be ready to escalate as delays arise
  • Give the auditee deadlines for deliverables
    • Clear
    • Attainable
  • Have an engagement letter for the DA
Delays from Auditee

- Auditee lawyers up
  - Have your legal counsel prepared ahead of time when:
    - You do not already have the data in hand before the DA
    - The auditee has a history of using legal resistance
    - Any time you are asking for data outside of another audit
      - Have an engagement letter for the DA
Data MoU / Engagement Letter

**Objective:** Documented leadership-level agreement for data access

**Achieved by mutual commitment to maintain CIA Triad**

- Testing queries before production access
- Establishing operational windows for queries
- Restricting query access to select trained personnel
- Ensuring account access is “read-only”
- Tune queries to minimize impact on system performance
- Establish protocols to secure data in custody of audit
- Establish protocols for data retention and destruction

CIA Triad

[Image of CIA Triad: Confidentiality, Integrity, Availability]
The DA team is pulled in different directions

• Make sure that any Auditors on the DA team are free to work on DA
  • It can be difficult for an auditor that may have been the lead on an audit for several years to be moved to a different project
    • If needed, create time in the DA timeline to allow for the Auditors to answer questions for the new audit lead from their old audit
    • Set clear guidelines for this process
    • Get mid and upper level management buy in of these guidelines
    • “Context shifting” impedes DA development – scheduling compromises
  • This can be especially difficult when borrowing an Auditor

• Clear communication is the key
You have a team and a scoped DA... what else?

- A couple of additional items to think about
  - You need a DA life cycle model
  - What are the next steps after the DA is complete
DA Model

• Agile
• Agile
• Agile
Agile DA Model

**Agile Software Development (ASD)** is an approach to software development under which requirements and solutions *evolve* through the collaborative effort of the team.

ASD is in contrast to a more traditional **Waterfall Software Development** model that is linear, and does not easily allow for the discovery of new information to influence the way the software is written.
Agile DA Model

- **Agile Software Development**
  - Does not mean scope creep
  - ASD allows for both data and programming directions to shift if it becomes apparent that there is a need
    - More efficiently achieves the desired project results.

- For example:
  - Shifting programming languages if the one the project was started in turns out to be problematic in using a Machine Learning module that best fits your data
  - Or finding alternate data sources for parts of your data needs
DA Model: CRISP-DM

**CRoss Industry Standard Process for Data Mining (v 1.0)**

- Non-Platform and Non-tool specific data mining methodology
  - Methodology can be applied to any software
- Developed in 1999-2000 and still in wide use
  - Industry standard data mining model
- DA Life-Cycle driven
- Agile
DA is the Beginning, Not the End

- Use the DA results to start audits
- Vet all results before reporting
What’s Next? The State Auditor DA Group

An open forum, meeting regularly, for all employees who work for a State Auditor to exchange information about what works well with DA and what doesn't

- Ideas
- Audits
- Software
- Getting past roadblocks
- Avoid duplicating efforts
- National data
- Multi state results
  - getting the Feds to notice
Questions?

Contact Dan Boulden (dan_boulden@ncauditor.net) for more information about the State Auditor Data Analytics group