Data Analytics on a Shoe String Budget

Presented By:

The Mississippi Office of the State Auditor

Stephen Houston, Director of Data Services
Sam Atkinson, Director of Special Projects

Shad White

Mississippi State Auditor

So your boss just told you he wants to "Analyze Big Data" and has tasked you with getting the job done...

© MARK ANDERSON, WWW.ANDERTOONS.COM

"According to all the big data we've gathered, our discussions about big data are up 72% this year alone."

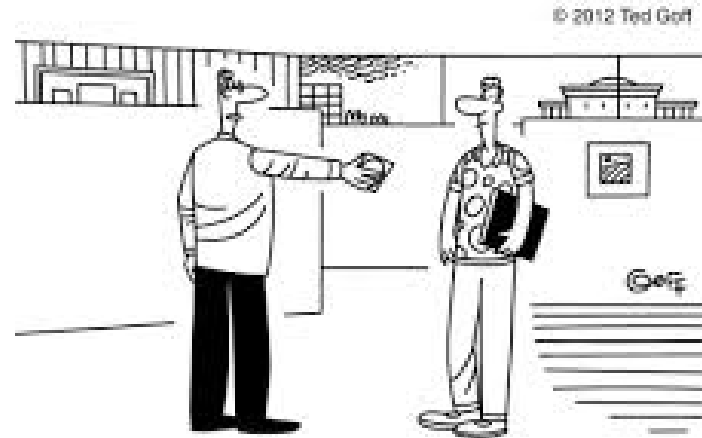**Shad White**

Mississippi State Auditor

Shad White

Mississippi State Auditor

First Thoughts

▸ Your initial thoughts and concerns are probably related to any or all of the following:

- Budget
- Staffing
- Staff Expertise
- IT Infrastructure (Hardware)
- Software Requirements
- Data Security
- Implementation Process
- Time Required

© 2012 Ted Goff

"Here's a list of 100,000 warehouses full of data. I'd like you to condense them down to one meaningful warehouse."

Shad White

Mississippi State Auditor

What You Know

- Your Current Operating Budget
  - OSA is appropriated approximately $8.8 million from the State for 170 positions
  - OSA is allowed to bill counties and school districts up to $3.2 million to offset local government audit costs.
  - Total: Up to $12 million or (average) $70,588 per person

- Your Current IT Budget
  - Less than $250,000 per year (all equipment, personnel, software, training, etc.)

- *Maybe* your Personnel situation

Fearless, You Proceed!

> ▸ After just a little research and a few phone calls to your finance director, you know there are some simple, easy, and cost efficient ways to get started…You can use:

> ▸ Existing staff

> ▸ Existing IT infrastructure that you already have, or get creative with obtaining additional infrastructure

> ▸ Affordable, yet powerful software to do your analysis

> ▸ Industry standard security practices and simple methods to alleviate security concern

Shad White

Mississippi State Auditor

Existing Staff

‣ In your agency, you likely already have staff that excel at the skills you will need to quickly begin analyzing big data.

‣ **Policy people.** They understand governmental policy, laws, regulations, and contracts, and can quickly adapt and learn the information surrounding the data you are analyzing.

‣ **IT people.** The IT staff at your agency will know how to set up or obtain an environment where you can store and work on the data you will need. They will also have the knowledge to make sure you are meeting security concerns.

‣ **IT people part 2.** Look for those who have solid working knowledge of database administration and know the ins and outs of how to work with data sets and clean them, organize them, join them, and query them.

Shad White

Mississippi State Auditor

IT Infrastructure

▸ Create a secure environment to work on and store your data ***using existing infrastructure***

▸ If using existing infrastructure is not an option, ***get creative*** with obtaining the equipment and storage you need—surplus property, other agencies transfers, etc.

▸ See if your IT department has old equipment that they have retired, but not yet disposed of laying around. If so, you may be able to ***repurpose it*** to create the infrastructure you need.

▸ Many government entities will ***transfer equipment*** to others at no cost when they upgrade/retire their infrastructure. Keep an eye out for equipment such as servers and storage arrays that may become available.

Shad White

Mississippi State Auditor

Now for the Software...

‣ Powerful Analysis Software can also be *free/low cost*:

‣ Free software, like **R Studio**
  ▪ While it has a high learning curve and is not as intuitive as other software, it can do most of the jobs you will need for analyzing your data.

‣ Affordable software (Depending on which products you get and how you set them up)
  ▪ **Tableau**
  ▪ **ACL GRC**
  ▪ **IDEA**
  ▪ **SPSS**
  ▪ **SAS**

Using any one of these or a combination of these will give you the ability to analyze your data in a meaningful way.

Shad White

Mississippi State Auditor

Or, Alternatively...

- Although OSA did not go this route, many small entities end up using a Hadoop solution instead of an RDBMS (Relational Database Management System) solution

- What is Hadoop?

- How does it work?

- How does the cost compare to a traditional RDBMS?
  - Hint: It is generally at least one-third the price, and many times even less

- Other Alternatives?
  - Other low cost alternatives
  - Contracts with the private sector
  - University interns/partnerships

Shad White

Mississippi State Auditor

Securing Your Data

‣ So you have signed your confidentiality agreements to get your hands on the data…

‣ IT staff should make sure that your analytics environment follows industry standards and privacy laws. This may include:

- Written security policies including incident response measures
- Full disk encryption
- Firewall rules
- Physical security
- Logging and monitoring
- Password guidelines, User Access Control, and auto time-outs

‣ Your IT staff can also recommend simple yet efficient methods for reducing your security risk. These may include:

- Setting up your analytics environment to be offline and not accessible via the internet or outside your office

Shad White

Mississippi State Auditor

Accessing Your Data

- Batch?

- Interactive?

- Realtime?

- Remote access allowed?

- Accessing on a portable device?

- How long do you keep your data?

Shad White

Mississippi State Auditor

Using Your Data

- Complexity
  - Federal Laws
  - Federal Regulations and Policy
  - State Laws
  - State Regulations and Policy
  - Government Contracts
  - Public Private Contracts

- Scope
  - Analysis time and deadlines
  - Size of data sets
  - Available time, system, and personnel

- Refining Analysis
  - Look for anomalies
  - Look for unusual patterns
  - Don't take data results at face-value

- Reporting

Shad White

Mississippi State Auditor

Putting The Pieces Together

The Mississippi Office of the State Auditor decided to begin analyzing big data in 2017

- ## The Team:
  - Stephen Houston – Software developer/analyst and database administrator
  - Samantha Atkinson – Policy expert

- ## The Hardware:
  - Retired storage arrays and servers to create our own infrastructure separate from our production environment
  - This made it easier to secure and allocate all resources on these machines for data analytics

Shad White

Mississippi State Auditor

Putting The Pieces Together

- ## **The Software:**

  - We purchased Tableau at an annual cost around $1,000/per license. We also already had ACL GRC which has a similar cost as Tableau.

- ## **The Data:**

  - Within one month, we began to obtain and analyze our data.
  - Within 8 months, we had identified $650,000 in irrefutable improper Medicaid Managed Care claims, which have been recovered.

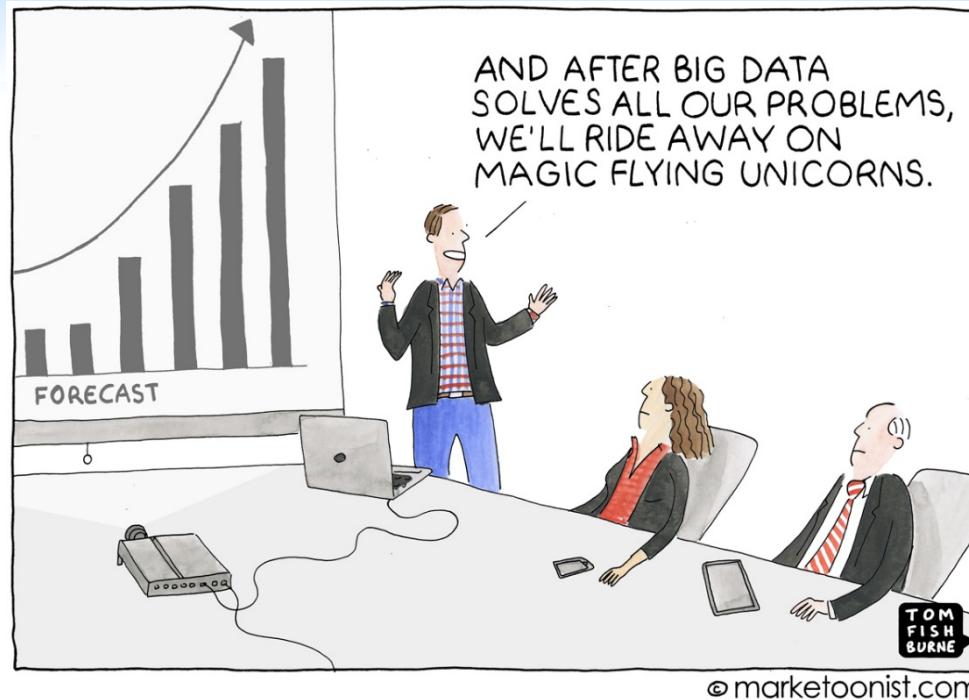- ## **The Projects Since Then:**

  - Medicaid
  - Education Expenditure
  - Student Browsing History
  - Government Accounting analysis (payroll, expenditures, etc.)

Shad White

Mississippi State Auditor

Thank You

Mississippi Office of the State Auditor

501 North West Street, Suite 801A
Jackson, MS 39201
Ph: 601-576-2800
800-321-1275 (statewide)

Website: www.osa.ms.gov

Shad White

Mississippi State Auditor