SAS – A Contemporary Analytics Environment

Adrian Jones
Advisory Business Solutions Manager, SAS
Two Mindsets

**ABUNDANT MINDSET**
- Technology empowered
- Discovery-centric
- Focus on value
- Everything is permitted unless it is forbidden

**SCARCITY MINDSET**
- Technology constrained
- Process-centric
- Focus on cost control
- Everything is forbidden unless it is permitted

Experimentation + Innovation
Productionization
Innovation & Experimentation = Market Leadership…

Innovation should not be constrained by the data you already have or can “afford” to use!

- Experimentation should enable Innovation
- Experimentation needs data from within AND outside your organization – in fact there should be NOTHING that is off limits respecting privacy of course
- Innovation, hopefully, leads to new services and new approaches that can be moved into production in a targeted and governed manner to ensure compliance amongst other things
The Big Data Lab - Unlocking possibilities

No such thing as a bad outcome just the result of testing an idea!
The Big Data Lab - Unlocking possibilities at the lowest possible cost with accessible skills

**Hardware**
- Commodity HW
- Low Entry Point
- Easy to expand
- Cheap Storage/Disk

**Software**
- Scalable Foundation
- Data Ingestion
- Data Transformation
- Data Visualization
- Broad Array of Analytics

**Skills**
- Programming
- Data Manipulation
- Analytics
  = Data Scientist or just an inquisitive mind?

Mandate is to come up with ideas for transformative disruptive change not just incremental improvements
The Big Data Lab is driving the adoption of Hadoop at MANY organizations
Hadoop Adoption - What are they doing right now?

- Improve asset utilization using machine data
- Risk Assessment & Fraud Detection
- Customer behavior & Sentiment
- Clinical Trials and Genomic Research
- Product, Content Recommendations
Looking for **efficiency** and **repeatability** to produce something of the highest value at the lowest cost possible.
Next Generation Analytic Platform - Need to support ALL parts of the Analytics Lifecycle

- Repeatable steps for obtaining the most value from all styles of analytics
- Leading companies often recognise the need for a process, standardize it and then continuously focus on speeding up the cycle
- A problem need not touch every part of this cycle
- Big Data is already having an impact on the whole of the analytics lifecycle
The new wave - Can Hadoop play a major role going forwards in the factory?

- The EDW and Data Marts are not going away
- The role of the EDW might change
- Ad-hoc BI and Analytics does not need an EDW designed for transactions!
- Is Hadoop the technology to free us from the shackles?
- Skills in Hadoop technologies is a big issue
Many possible roles for Hadoop…
Our strategy is to enable the entire lifecycle around HADOOP

...and now we are delivering solutions on top of this such as SAS High Performance Anti-Money Laundering, SAS High Performance Risk, SAS High Performance Marketing Optimization etc.
Our strategy is to enable the entire lifecycle around HADOOP (and RDBMS if that is your choice)

Visual Analytics: Also available to run directly on Oracle Exadata, Teradata, Pivotal (Greenplum) or on a standard server reading data from almost anywhere.

Scoring Accelerator: Also available for SAP HANA, Oracle Exadata, Teradata, Pivotal (Greenplum), IBM DB2, Netezza (IBM Pure Data for Analytics)

Code Accelerator: Also available for Teradata, Pivotal (Greenplum)

Done using the Build Model Tools and other checks

SAS Data Loader for Hadoop (Includes the DQ Accelerator for Hadoop and the Code Accelerator for Hadoop)

Support for most major RDBMS systems

Visual Analytics and Visual Statistics: Also available to run directly on Oracle Exadata, Teradata, Pivotal (Greenplum) or on a standard server reading data from almost anywhere.

Also available for Oracle Exadata, Teradata, Pivotal (Greenplum)

Done using either the Data Preparation, Data Exploration or Build Model Tools
Hadoop + SAS – a contemporary analytics platform

- Hadoop is **scalable** and **cheap** (storage & processing)
- Hadoop is rapidly **evolving** and **maturing**
- Hadoop is rapidly becoming **mainstream** technology
- Major marketplace interest, activity and **spend**
- Increasing adoption rates by **your competitors**
- Deals with two needs – **cost reduction** and **innovation**
- The technologies already **exist** to help from SAS!
The overwhelming *majority* of organizations have neither a finely honed analytical capability nor a detailed plan to develop one.

Source: Competing on Analytics