From Theory to Data Product

Applying Data Science Methods to Effect Business Change

Strata Data Conference NY 2018 – September 11
Components of an Advanced Analytics Project
Are you talking to the right people?
Advanced Analytics Entry Points

- User Driven
- Management Driven
- IT
- Business
Build Brilliant Together

User Driven

Management Driven

The Technology Directive

Business
Results were questioned and not aligned with current processes
“Our vendor is really confident with this technology!”
“We do need Data Governance and we’ll do that once we’ve shown value”
“I’m not sure who will look at it but it will be interesting”
“It’s not that easy”
“The data is pretty self-evident”
“We didn’t schedule time with [the SMEs] for that”
The Checklist:
Build Brilliant Together

The Technology Directive

User Driven

Management Driven

Business
The Field of Dreams

The Technology Directive

User Driven

Management Driven

Business
The right business people were not identified project resources
No executive support to move forward
“We have easy access to the data”
“If we build it, they will come”
“I can’t tell you what that means, you need to talk to...”
“I can’t get time with [SME] to...”
“What happens if this information gets out?”
“Why wasn’t I involved?”
The Checklist:

- Identify business questions
- Identify required data
The Field of Dreams

The Technology Directive

User Driven

Management Driven

Business
End users saw no value in, or use of, the results
“We don’t need to talk to the end users”
“It doesn’t matter what our current metrics are”
“We only use standard industry terms”
“We are just implementing what the business tells us”
“Are we sure this is what the users want?”
The Checklist:

- Identify the business SMEs
- Have SMEs identify business drivers & focus on them
- Identify business questions using drivers as a guide
The Field of Dreams

The Technology Directive

User Driven

Management Driven

The Ambitious Executive

Business
The Field of Dreams
User Driven
The Smart Competitor
Business

The Technology Directive
Management Driven
The Ambitious Executive
“We don’t know what we don’t know”
“The data is self-evident”
“That’s third party software. We don’t have a data dictionary”
“Doesn’t seem like they are very interested”
The Checklist:

- Identify the business SMEs
- Have SMEs identify business drivers & focus on them
- Identify business questions using drivers as a guide
- Prioritize actionable business question
- Identify required data
- Identify and engage data stewards
- Choose your resources
- Create a data dictionary
The Field of Dreams

The Technology Directive

The Ambitious Executive

The Smart Competitor

User Driven

Management Driven

Business
The Checklist:

- Identify the business SMEs
- Have SMEs identify business drivers & focus on them
- Identify business questions using drivers as a guide
- Prioritize actionable business question
- Identify required data
- Identify and engage data stewards
- Choose your resources
- Create a data dictionary
- Check back against business drivers throughout
The Technology Directive

Management Driven

Project Start
- Identify Business SMEs
- Identify Business Drivers

Midstream
- Discuss business SME resource gaps with PM/Project Sponsor
- Gather results to date
- Engage business SMEs in workshop to focus on actionable insights
The Field of Dreams

= User Driven

- Identify Business SMEs
- Identify Business Drivers
- Talk to business counterpart to garner support and participation
- Identify proper data stewards
- Engage business SMEs in workshop to focus on actionable insights
The Ambitious Executive

Project Start

Identify Actionable Business Question

Midstream

Talk to end business users to best understand process and fit

Identify Business Drivers

Management Driven

Business
The Smart Competitor

User Driven

Project Start
- Engage C-Suite early

Midstream
- Check in with business drivers and strategy
- Realign if necessary
Advanced Analytics Entry Points

The Checklist:
- Identify the business SMEs
- Have SMEs identify business drivers & focus on them
- Identify business questions using drivers as a guide
- Prioritize actionable business question
- Identify required data
- Identify and engage data stewards
- Choose your resources
- Create a data dictionary
- Check back against business drivers throughout
Group Exercise
Are you asking the right questions?
Here’s what we’re trying to avoid
Components of an Advanced Analytics Project
When answered, a specific action can be taken which provides a measurable result in line with strategic business objectives.
Our Process

Data Driven Decision Making

Discovery

Map Influence Points
Risk: Value: Readiness
Experimental Modelling
Proof of Value

Pilot
Enterprise Rollout
Enterprise AI

64
Discovery

Driver Definition Workshop

Business Value Workshop

Project Kickoff
Discovery

Driver Definition Workshop  Business Value Workshop  Project Kickoff
Business Value Workshop

... the first hour...
Non-Starters

- Can we show value with tool X?
- Data science helps businesses – how can it help ours?
- How can we establish KPIs to more clearly indicate when high-level actions should be taken?
We need to use deep learning to do some stuff...

Produce an algorithm that can be used, season to season, to set prices when new products become available in the market from us or competitors

Create a full list of subscribers and their respective churn probabilities, indicating probability to churn within 4 weeks
Requires Refinement

- Who are the biggest users of healthcare in ...?
- How can we save $ on our preventive maintenance?
- When will customers leave?
- What products / features should we build in?
Business Value Workshop

... the first hour...

... hour 2 and beyond ...
Refining

Brainstorm
(Workshop!)

Evaluate
(Map!)

Refine
(Key Questions!)
Scenario 1: Telecommunications
In the ‘business value workshop’...

**STARTED HERE:** “We want a machine learning model that predicts customer churn four weeks in advance of the event, with an accuracy of at least 80% and lift curves graphed to demonstrate why you picked this model.”
“We want to know when our customers will leave.”

“Well, it doesn’t always. Only some of the customers leaving actually matters.”
“High-value customers matter... when they leave, they take a lot of business with them.”

“Um... Predicted lifetime value = $x/time unit/y years”

“Well, we’ll make sure customer service knows so they’re careful with people who call in with a complaint. And maybe we’ll offer them something to make staying more appealing.”

“A special offer... we’d need at least two months to put something like that out. And--how long do we need to change customer service approaches before someone leaves?”

Can we predict which of our high-value customers will leave in two months?
Candidate Business Question

Can we predict which of our high-value customers will leave in two months?
Scenario 2:

Healthcare
In the business value workshop

STARTED WITH: Which of our company’s three options is the most effective treatment approach for a given patient?
Why? Who? What? Where? Do we have it?
Candidate Business Question

How do you predict the optimal treatment for any given patient before treatment is prescribed?
Refining

Brainstorm
(Workshop!)

Evaluate
(Map!)

Refine
(Key Questions!)
Let’s dive in to one example

1. How do you predict the optimal treatment for any given patient before treatment is prescribed?

2. What other products should we develop as part of our existing product line? (predict market needs)

3. Can we compile data from our healthcare providers to understand optimal practices in installing our products based on outcomes (recovery time)?

4. Can we determine poor performing products before they leave the factory?
Prioritization Criteria

• Value
• Feasibility
• Alignment
Value
Feasibility
Alignment
Prioritization for our scenario
Valuable Business Question: When answered, a specific action can be taken which provides a measurable result in line with strategic business objectives.
Group Exercise
Are you following a solid process?
Components of an Advanced Analytics Project
What is our definition of Agile?

• **Adapted** for data driven decision making projects
• An iterative and incremental approach to project management and delivery
• Allows project teams to adapt quickly to changing requirements and/or new insights
Managing Uncertainty

• Experiment design is focused on constant progress toward the stated goal
• By their nature, experiments may not yield expected results
• Project strategy should be to adapt subsequent experiments, based on findings, to continue progressing toward the goal
• Using this approach mitigates risk
  • Key activities early in the project validate feasibility of moving forward
  • Goal is to fail fast – reassess throughout to avoid unproductive activity
Our Process

Map Influence Points  
Risk: Value: Readiness  
Experimental Modelling  
Proof Of Value  
Pilot  
Enterprise Rollout  
Enterprise AI

Data Driven Decision Making
Proof Of Value Activities

Map Influence Points

- Influence Point Mapping
- Influence Path Identification

Risk Value Readiness Assessment

- Per Path:
  - Data Asset Assessment
  - Risk Identification
  - Business Value Assessment
  - Prioritization Map

Experimental Modeling

- Experiment Design
- Experiment Implementation
- Experiment Testing
- Analysis
Enterprise AI Activities

Pilot Sprints 1-n
- Refine Model
- Implement Infrastructure
- Deploy Model

Validation
- Validate Model Performance
- Validate Insight(s)
- Validate Actions
- Report on ROI Results

Enterprise Rollout Sprints n+1 - z
- Scale Implementation
- Define Ongoing Success Criteria
- Check for Model Drift
- Ongoing DevOps

Operational Maintenance
- Measure Ongoing Success
- Evaluate and Plan

Build Brilliant Together
Data Science Elements of a Project

• Focused on data
  • Ex. Feature analysis

• Focused on testing results of a model
  • Ex. Cross validation, lift, gain, etc.

• Focused on testing the insight
  • A/B Testing, focus groups, etc.

• Focused on data product
  • Checking for model drift, model maintenance
  • Ex. Recommendations engine, etc.
## Roles and Involvement – POV

<table>
<thead>
<tr>
<th>Activities</th>
<th>Architect</th>
<th>Data Science Strategist</th>
<th>Data Quality Analyst</th>
<th>Data Science Team</th>
<th>Designer / Developer</th>
<th>Product Owner / Subject Matter Expert</th>
<th>Scrum Master / Project Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Map Influence Points</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Risk Value Readiness Assessment</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Experimental Modeling</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
## Roles and Involvement – Enterprise AI

<table>
<thead>
<tr>
<th>Activities</th>
<th>Architect</th>
<th>Data Science Strategist</th>
<th>Data Quality Analyst</th>
<th>Data Science Team</th>
<th>Designer / Developer</th>
<th>Product Owner / Subject Matter Expert</th>
<th>Scrum Master / Project Manager</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot Sprints 1-n</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Validation</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Enterprise Rollout</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Operational Maintenance</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
Group Exercise