How to establish and sustain a data insights capability in large corporations.
Introduction

A little about me ...
Why this presentation

*What I am hoping you will get out of today.*

This is a reflection on my experiences, at a practical level, in:

- Defining and establishing data insights teams
- Facing challenges and hurdles on the journey
- Finding options to overcome the challenges
- Discovering what makes a good data insights team (hint its not all about data scientists)
- Structuring and staffing for success

To quote GoT as a summary of my journey ...

*The night is dark and full of terrors*
Building data science teams is easy right? Plenty of reference materials available. Articles Listed on DataCamp web.

1. So you want to build a data science team? - VentureBeat
2. Tips for Building a Data Science Capability - Booz Allen Hamilton
3. Three Essential Components of a Successful Data Science Team - KDNuggets
4. Building Data Science Teams - DJ Patil (O'Reilly)
5. 9 Tips for Building Data Science Teams - Georgian Partners
6. Doing Data Science Right - FirstRound
7. Don't look for unicorns, build a data science team - CIO
8. 5 Tips for Building a Great Data Science Team - Forbes
9. Building a Data Science Team - A Moneyball Approach - Vijay Bhat
10. How To Win The Hiring War for Data Scientists - FastCompany
11. How to Consistently Hire Remarkable Data Scientists - FirstRound
12. Using R packages and education to scale Data Science at Airbnb - Ricardo Bion (Data Science @Airbnb)
13. 3 changes that will strengthen your data science team - VentureBeat
14. The Data Science Industry: Who Does What (Infographic) - DataCamp
15. Seven traits of successful Data Science teams - Teradata
16. 4 Online Data Science Options for Your Team - KDNuggets

There is no substitute for just getting on with it ... BUT before you do, ask yourselves these simple questions...
Before we begin ...

Do I really need a data insights / science capability?

Is my company ready for a new capability?

Can I make a difference to my company’s operations?
But, we already have a BI analytics team...

Extracting data from core operational systems for reporting
Create an abstracted model to store the data.
Transform and restructure that data to conform to the model.
Write queries and reports for business users.
Take extracts of the data for external parties (regulators, etc)
Restructure the data into user friendly departmental/functional views.
Provide power users with tools to access, query and visualise the data.
Rinse and repeat ...

Works just fine for many organisations today ... but you are locked into this method of information development and analysis.
Organisations are now demanding broader use of information to meet their corporate objectives

<table>
<thead>
<tr>
<th>Increase Time to Insight</th>
<th>Processing &amp; analysing huge volumes of disparate data in real time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifying &amp; Collecting New Data</td>
<td>Leveraging open data and other sources of enrichment</td>
</tr>
<tr>
<td>Integrating Disparate Data Types</td>
<td>Combining video, audio, logs, text, social and structured data for analysis</td>
</tr>
<tr>
<td>Moving Hindsight to Insight</td>
<td>Complimenting BI &amp; Reporting with techniques to generate valuable insights</td>
</tr>
<tr>
<td>Moving Foresight to Action</td>
<td>Generate actionable results through predictive modelling and effective visualisation</td>
</tr>
<tr>
<td>Developing new hypotheses</td>
<td>Asking questions not previously possible to learn things you previously did not know</td>
</tr>
<tr>
<td>Test, Learn &amp; Adapt loop</td>
<td>Learn and Adapt to further optimise outcomes and continuously improve.</td>
</tr>
</tbody>
</table>

Represents skills and experiences not normally found within typical BI analytics teams.
Ok. So you **do** need a data insights capability

*What are the potential hazards & roadblocks?*

**Cultural**
- Challenges to authority
- Ownership issues
- Not data driven
- Low perceived data value
- Sharing & openness

**Organisational**
- Location within company
- How to staff the new function
- Central vs distributed
- What positions are needed
- How to govern / measure

**Technical**
- Skills availability
- Internal technology too slow
- Too many tools to choose
- Architecture for a new world
- Hadoop → Spark → Cloud → ?
“Data-driven organizations are three times more likely to report significant improvement in decision making. But only 1/3 of 1,135 senior executives surveyed consider their organizations highly data-driven.”

Source: PwC Global Data and Analytics Survey
Encourage a data-driven culture

_Cloudera recommend these 6 core principles of an enterprise data architecture_

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data is a Shared Asset</td>
<td>Information is still Power in many organisations. This seriously limits the desire to share.</td>
</tr>
<tr>
<td>Readily Available for Use</td>
<td>If users cannot consume the data, it holds no value. Provide interfaces and tools for this.</td>
</tr>
<tr>
<td>Security and Access</td>
<td>Unified data security, increases confidence.</td>
</tr>
<tr>
<td>Provide a Common Vocabulary</td>
<td>A shared asset needs to be understood. Reduce ambiguity and increase trust.</td>
</tr>
<tr>
<td>Stewardship and Curation</td>
<td>Lack of ownership kills many data programs. Without curation, self-service is frustrating.</td>
</tr>
<tr>
<td>Eliminate Movement &amp; Copying</td>
<td>Ensure freshness of data, reducing cost and increase data agility by keeping data close.</td>
</tr>
</tbody>
</table>
Communicate to dispel FUD

New team, new ideas, new tools and new methods means many in your organisation are unfamiliar with what you do .. which can lead to fear.

What impact will you have on other’s jobs, reward, careers?

Regular and open communication is critical in the success of any new data insights team.

Involve and collaborate broadly and regularly.
“Speed to insight—and being able to determine what needs to go to market, and how efficiently we can get it there—allows us to stay nimble.”

Source: Pawan Divakarla, Data & Analytics Leader, Progressive
So what makes a good insights team

**Governance (Cop)**
Applies the *Laws* of ownership, quality, lineage and *Responsible road use* of the data. *Stopping* others from dangerous misuse of data, including *Theft*.

**Data Wrangler (Cowboy)**
Expert at *Roping* in data of different types across a broad *Landscape*. *Corrals* data into usable *Pens* for *Keeping* and *Blending* with other data as needed.

**Developers (Construction)**
*Building* and *Maintaining* the Data Pens and Stables, the Data Science Machines and Bikes, and the Charting, Navigation and Visualisation tools.

**Visualisation (Sailor)**
Creates *Charts* and *Maps* of the data, making it easy for users to *Navigate* the information to get clear *Direction*.

**Data Scientist (Biker)**
A *Hardcore Gang* making *Unorthodox* use of statistical algorithms to *travel* further down the road at *breakneck speeds*.

**Leader (Chief)**
Removing organisational impediments, sharing the *Pipe of Peace* with other departments, allowing *Data Warriors* to *Roam* and *Hunt* for data freely.
Structure for actionable insights

Centralised <-> Decentralised or something in-between?

Distributed
Capabilities developed within each business unit with no central control.

Business-Led
Predominantly business-led with some support being provided centrally.

Centre-Led
Mainly centralised with business capability retained for specific use cases only.

Centralised
Capabilities are built at an enterprise level within a centralised function.

Executive Collaboration, CxO Sponsorship, Strategic Alignment, Governance & Prioritisation

Head of Business Operations

Strategy & Context

Head of Enterprise Insights

Insight Generation

CORE SQUAD

Enterprise Insights

Use Case Discovery
User Training

Best Practices & Methodologies
Communications
Drive Analytical Outcomes

Data Services

DELIVERY SQUAD

Corporate Data
BI Reporting
Master Data

BI Architecture
New data sources
Metadata mgmt

Balance of Skills Across Organisation – Highly Collaborative – Talent Leverage

Common Tools, Methodologies & Practices

© COPYRIGHT INFOSTRATEGY 2017
Staffing

**CoE Model – leveraging internal and bring external resources**

Find data scientists (mix and match skills)
- SPSS Statistician
- Python skilled Data Engineer

Align and compliment business analytics skills (negotiated secondment / training)
- Marketing and Customer Analytics
- Business and Finance Performance Analysts
- Digital Analytics (clickstream and interaction)

Supporting skills
- Information / Data Analyst – Core Team
  - Stakeholder engagement (internal transfer)
- Analytic Insights Architect – Core Team
  - Research and identify suitable tech capabilities we can leverage
  - Run POCs on selected technologies
  - Support Insight Generation team with tools and training
- Governance (Data and Algo) – Core Team
  - Owner and Stewards in the business.
  - Build data catalog of data assets

Leverage partners to accelerate capability
**Insights operating model**

*Freedom, creativity, experimentation, test-and-learn*

- **Diagnostic & Descriptive**
  - Corporate Performance Management Reporting

- **Discovery**
  - Data Mining

- **Predictive & Prescriptive**
  - Strategic Planning

**Hindsight**

Your existing BI capability is critical for your reporting and analysis needs (operational BI dashboards etc.)

**Insight**

Your insights team needs to have the freedom to be creative, to drive innovation, come up with new ideas, suggestions and data products.

Feeding both sides of the analytic equation.

**Foresight**

Your operating unit data analysts / scientists are focused in improving current processes, creating new processes based on propensity and simulated outcomes.
Wrap up & summary

Key take-away reminders from this session

**Be Certain.** Be clear about what you want from your advanced analytics capability. Back yourself and your team.

**Be Brave.** There are many challenges. But understanding them and being prepared will lead to success.

**Be Creative.** This is a new world. There are no hard and fast answers.

**Be Outcome Focused.** Business-led and context aware.

**Be Communicative.** Break down the FUD barriers. Engage widely.

**Become a Data-Driven Culture**
Thank you

Open for Questions
## Structure for actionable insights

### Centralised <-> Decentralised or something in-between?

<table>
<thead>
<tr>
<th>Distributed</th>
<th>Business-Led</th>
<th>Centre-Led</th>
<th>Centralised</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capabilities developed within each business unit with no central control.</td>
<td>Predominantly business-led with some support being provided centrally.</td>
<td>Mainly centralised with business capability retained for specific use cases only.</td>
<td>Capabilities are built at an enterprise level within a centralised function.</td>
</tr>
<tr>
<td>✓ Highly responsive to the business need</td>
<td>✓ Business retains control of analytical activities and investments</td>
<td>✓ Central control of analytical activities with input from business</td>
<td>✓ Economies of scale lower cost and increase efficiencies</td>
</tr>
<tr>
<td>✓ Deep business context</td>
<td>✓ Capability still close to decision point</td>
<td>✓ Increased reuse through consolidated core capabilities</td>
<td>✓ Standardised tools, processes and methods</td>
</tr>
<tr>
<td>✓ Capability at decision point</td>
<td>✓ Power users remain connected and accountable to operating units</td>
<td>✓ Consistency of understanding business assumptions and analytical outputs</td>
<td>✓ Ability to deliver against group/enterprise outcomes</td>
</tr>
<tr>
<td>✓ Improved collaboration between business and central teams</td>
<td>✓ Improved collaboration between business and central teams</td>
<td>✓ Improved reporting processes</td>
<td>✓ Data treated as a strategic enterprise asset</td>
</tr>
</tbody>
</table>

- D Duplication of capability leads to higher cost
- D Reliance on individual staff members
- D Cross-function interaction and sharing of insights limited.
- D Challenging to maintain data standards and governance across the organisation.
- D Lack of common definitions and processes
- D Challenges in governing analytical standards and methods
- D Individual data solutions persist
- D More time spent in reconciling business-level and group outputs
- D Strong political leadership needed to traverse organisational structures
- D Negotiation with business units to provide resources / control
- D Significant change control and communications required
- D Strong demand management and prioritisation required
- D Isolation from business unit (context)
- D Back channel (shadow) requests can undermine success of a centralised approach.