What 300 L&D leaders have learned about building data fluency

Hugo Bowne-Anderson
@hugobowne

11/21/2019
Hugo Bowne-Anderson, data scientist at DataCamp

- Undergrad in sciences/humanities (double math major)
- PhD in Pure Mathematics (UNSW, Sydney)
- Applied math research in cell biology (Yale University, Max Planck Institute)
- Python curriculum engineer at DataCamp
- Host of DataFramed, the DataCamp podcast
Why data? Why now?
The value of data fluency
Best practices for data fluency
Biggest roadblocks to data transformations
Today’s topics of discussion

➔ Why data? Why now?
➔ The value of data fluency
➔ Best practices for data fluency
➔ Biggest roadblocks to data transformations
Why data?
Illustrations you can use, just copy/paste

Why data?

NETFLIX VS. BLOCKBUSTER (1998 - 2016)

Source: Disrupted or Disrupter: Which One Will You Be? (Cloud Technology Partners, 2017)
Illustrations you can use, just copy/paste
Why You Should Care About Data Science

➔ When so much data is informing decisions across so many industries, you need to have a basic understanding of the data ecosystem in order to be part of the conversation.

➔ On top of this, the industry that you work in is already seeing the impact of data analytics.
Definitions of Data Science

➔ “A data savvy, quantitatively minded coding literate problem solver” -- Cathy O’Neil, 2013

➔ “Data science doesn't just predict the future. It causes the future.”
-- Cathy O’Neil, 2018
Why now?
McKinsey & Co., *An Executive’s Guide to AI*
Data Science: why now?

McKinsey & Co., *An Executive’s Guide to AI*
A Minute on the Internet in 2019
Estimated data created on the internet in one minute

- 3.8m requests
- 347,222 scrolls
- 1m views
- 87,500 people on Twitter
- 4.8m GIFs served
- $996,956 spent online
- 390,030 apps downloaded
- 41.6m messages sent
- 1m logging in
- 4.5m videos watched
- 2.1m snaps taken
- 1.4m swipes
- 46,200 new posts
- 694,444 hours watched
- 188m emails sent

Sources: Lori Lewis & Officially Chad via Visual Capitalist
Poised to Disrupt All Industries
Stock market prediction
“Traders used to be first-class citizens of the financial world, but that’s not true anymore. Technologists are the priority now...” -- Robin Wigglesworth
“...with a little bit of training, you can accomplish quite a lot compared to the traditional approaches in this field.” -- Yves Hilpisch
Data Science in Finance

- Process Automation
- Security
- Underwriting and credit scoring
- Algorithmic trading
- Robo-advisory

- BNY Mellon
- PayPal
- OnDeck
- 2sigma
- Betterment

- Integrating robotic process automation into their banking system
- Identifying suspicious account behavior
- Underwriting model evaluates thousands of historical loans to improve future credit quality of portfolio
- Analyzing many data sources simultaneously to make better trading decisions
- Optimizing portfolio management and personalizing investment recommendations
Cancer patient survival probability
Disease diagnosis from imaging data (e.g. MRI, PET)
Google AI have developed an algorithm to detect diabetic retinopathy as well as experts. One of the leading causes of blindness globally. Around 400 million people have this disorder.
## Value of homes on Airbnb

<table>
<thead>
<tr>
<th>id_listing</th>
<th>host_location_city</th>
<th>avg-nightly-price</th>
<th>availability_next_180_days</th>
<th>1_year_revenue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>London</td>
<td>$120</td>
<td>50 nights</td>
<td>$</td>
</tr>
<tr>
<td>2</td>
<td>San Francisco</td>
<td>NULL</td>
<td>150 nights</td>
<td>$$</td>
</tr>
<tr>
<td>3</td>
<td>Tokyo</td>
<td>$55</td>
<td>NULL</td>
<td>$$$</td>
</tr>
<tr>
<td>4</td>
<td>New York</td>
<td>$100</td>
<td>90 nights</td>
<td>$$ $$ $$</td>
</tr>
</tbody>
</table>

Robert Chang, airbnb
Don’t fear automation
Humans in the loop
Humans in the loop

I got fired from work today

They replaced me with a human

Oh really? What happened?
Today’s topics of discussion

➔ Why data? Why now?
➔ The value of data fluency
➔ Best practices for data fluency
➔ Biggest roadblocks to data transformations
The value of data fluency

When compared to the last 5 years, would you rate your company’s current performance as better or significantly better in the following areas?

- Revenue growth: 80% (Mature DF competencies, N=150), 49% (Immature DF competencies, N=95)
- Market share: 70% (Mature DF competencies, N=150), 40% (Immature DF competencies, N=95)
- Profitability: 78% (Mature DF competencies, N=150), 41% (Immature DF competencies, N=95)
- Customer satisfaction: 79% (Mature DF competencies, N=150), 45% (Immature DF competencies, N=95)
- Employee satisfaction: 69% (Mature DF competencies, N=150), 26% (Immature DF competencies, N=95)
Describe your company's priority level when it comes to building/improving data fluency.
Describe your company's priority level when it comes to building/improving data fluency.

- Not a priority: 1% (Mature DF competencies, N=150), 7% (Immature DF competencies, N=95)
- Low priority, in strategic discussions: 3% (Mature DF competencies, N=150), 17% (Immature DF competencies, N=95)
- Moderate priority, some actions already taken: 42% (Mature DF competencies, N=150), 51% (Immature DF competencies, N=95)
- High priority, on-going corporate focus: 55% (Mature DF competencies, N=150), 25% (Immature DF competencies, N=95)
The manifold dangers of immature data fluencies

There are several dangers for companies with immature data fluency competencies that don’t take action to catch up to their peers:

- They may lose competitive advantage and undermine business strategies.
- They may fall behind in individual, team, and cross-team efficiency.
- They may miss out on essential automation opportunities.
- They may have a harder time tracking successes and failures.
- They may fall behind in innovation capabilities.
- They may miss out on fostering employee engagement, since employees who are actively learning are more productive and ask better questions.
The value of data fluency

What value can data fluency add to your company?

- Improving company performance: 37% (Mature DF competencies, N=150), 18% (Immature DF competencies, N=95)
- Improving individual productivity: 44% (Mature DF competencies, N=150), 15% (Immature DF competencies, N=95)
- Increasing individual skill levels: 35% (Mature DF competencies, N=150), 16% (Immature DF competencies, N=95)
- Fostering employee engagement: 31% (Mature DF competencies, N=150), 14% (Immature DF competencies, N=95)
- Improving innovation capabilities: 47% (Mature DF competencies, N=150), 18% (Immature DF competencies, N=95)
- Supporting leadership dev.: 38% (Mature DF competencies, N=150), 18% (Immature DF competencies, N=95)
- Automating processes: 38% (Mature DF competencies, N=150), 24% (Immature DF competencies, N=95)
- Supporting agility: 40% (Mature DF competencies, N=150), 19% (Immature DF competencies, N=95)
- Providing competitive advantage: 42% (Mature DF competencies, N=150), 16% (Immature DF competencies, N=95)
Those who have already invested see much more added value in providing a competitive advantage and improving innovation.
Why data? Why now?
The value of data fluency
Best practices for data fluency
Biggest roadblocks to data transformations
Jacqueline Nolis breaks data science into 3 components (Episode 28, DataFramed):

1. Descriptive analytics (Business Intelligence)
2. Predictive analytics (Machine Learning)
3. Prescriptive Analytics (Decision Science)
1. **Business Intelligence (descriptive analytics)**

- Taking data company already has
- Getting that data to the right people
- In form of dashboards, reports, emails
2. Machine Learning (predictive analytics)

➔ Put models continuously into production
➔ E.g. LTV at airbnb
Getting Better at Machine Learning
Moving Beyond `model.fit(X, y)`

➔ **Problem Definition**: why thinking hard about your problem is crucial
➔ **Data Collection**: why setting up your `{X, y}` right is half of the job done
➔ **Model Building**: how to debug your model when it does not perform well
➔ **Productionization**: what “putting model into production” really means
➔ **Feedback Loop**: how unintended feedback loop can affect your system
Types Of Data Science

3. Decision Making (prescriptive analytics)

➔ Take the insight discovered from the data science work
➔ Use it to help company decision making
➔ E.g. what do you if your data science work tells you a particular type of customer will churn?
Rogati’s AI Hierarchy of Needs

The AI Hierarchy of Needs, Monica Rogati
Foundational skills are key

To what extent are the following data skills important to your company?

- **Foundational data skills**: 72% already invested, 58% future investment
- **BI and dashboards (descriptive analytics)**: 60% already invested, 53% future investment
- **Machine learning (predictive analytics)**: 55% already invested, 46% future investment
- **Decision science (prescriptive analytics)**: 58% already invested, 60% future investment
Taras Gorishnyy identified the necessary moving parts (Episode 30, DataFramed):

➔ Executive support
➔ Analytics vision
➔ Build data foundations
➔ Distribution of skills & establish data culture
➔ Establish impact of analytics early on in the process
Best practices for data fluency

To what extent has your company taken (or plans to take) the following actions to become data fluent?

<table>
<thead>
<tr>
<th>Action</th>
<th>Mature DF competencies, N=150</th>
<th>Immature DF competencies, N=95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a vision for analytics aligned with company strategy</td>
<td>44%</td>
<td>6%</td>
</tr>
<tr>
<td>Establish strong C-level support</td>
<td>35%</td>
<td>7%</td>
</tr>
<tr>
<td>Extract value early from at least several use cases</td>
<td>32%</td>
<td>12%</td>
</tr>
<tr>
<td>Implement process redesign and culture change</td>
<td>39%</td>
<td>7%</td>
</tr>
<tr>
<td>Build data infrastructure to provide integrated, trusted, and timely data</td>
<td>43%</td>
<td>20%</td>
</tr>
</tbody>
</table>
Why data? Why now?
The value of data fluency
Best practices for data fluency
Biggest roadblocks to data transformations
Challenges in your data transformation

To what extent do the following business challenges prevent your company from building/improving data fluency?

- Lack of digital strategy: 18% (already invested), 34% (future investment)
- Complexity of core business operations: 30% (already invested), 39% (future investment)
- Lack of C-level support: 18% (already invested), 36% (future investment)
- Difficulty hiring top talent: 29% (already invested), 46% (future investment)
- Instability of business plans: 20% (already invested), 34% (future investment)
# How to build data skills

### What actions has your company taken to build data skills?

<table>
<thead>
<tr>
<th>Action</th>
<th>Mature DF competencies, N=150</th>
<th>Immature DF competencies, N=95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Established a data university/Center of Excellence</td>
<td></td>
<td>37%</td>
</tr>
<tr>
<td>Provide on-site/in-person training</td>
<td></td>
<td>60%</td>
</tr>
<tr>
<td>Provided access to online learning platforms</td>
<td></td>
<td>64%</td>
</tr>
<tr>
<td>Launched internal campaigns to promote/incentivize learning</td>
<td></td>
<td>51%</td>
</tr>
<tr>
<td>Recruited employees with existing, relevant skills in data fluency</td>
<td></td>
<td>49%</td>
</tr>
<tr>
<td>Sourced solutions from a vendor to bridge gaps in data fluency</td>
<td></td>
<td>48%</td>
</tr>
</tbody>
</table>

Legend:
- Tape: Mature DF competencies, N=150
- Fill: Immature DF competencies, N=95
Thank you!

Hugo Bowne-Anderson
@hugobowne
@datacamp