Data 101:
How to use your data science team: Becoming a data-driven organization

Yael Garten
LinkedIn

September 29, 2015

http://www.linkedin.com/in/yaelgarten, @yaelgarten
Every organization, every product, every decision data driven, or at least data informed.
To create a data driven organization

1. Consciously map how you use data in each phase of the product lifecycle

2. Treat data as a first class citizen

3. Create a culture which expects decisions are informed by data
This talk will be a success if we:

1. Review the steps of data driven product innovation

2. Understand what is needed to best enable fostering of (or transformation into) a data-driven organization: culture, process, and tools.
Data driven product innovation framework:
Use data to measure, understand, and improve the product:

<table>
<thead>
<tr>
<th>Build</th>
<th>Measure</th>
<th>Track</th>
<th>Ship</th>
<th>Tweak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actionable insights lead</td>
<td>Success Metric Definition</td>
<td>Tracking Instrumentation spec</td>
<td>Experimentation Setup and Analysis</td>
<td>Post-Launch Analysis</td>
</tr>
<tr>
<td>to product feature</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If data scientist is not involved until this stage, it may be too late.
1. Opportunity sizing: how big or important is the problem?

2. Use data to predict successful product initiatives:
   - Show news articles
   - Suggest new connections
   - Suggest following active content creators
Hypothesis: Following active sources leads to improved user experience with LinkedIn Feed

Success Metric – progression of definition:

- Total clicks on Follow
- # clicks / #impressions of Follow suggestions
- % Feed Inventory created by new followees
- Downstream sustained engagement with items created by these followees
  - What is engaging? # of clicks? Time spent? # Shares? Combo?
Need accurate reliable standardized data logging to enable metric computation.

**Metric** = *Downstream engagement with items created by these followees*

Must enable attributing future clicks on feed items to that campaign as a source for the Follow.

```
FeedActivityClick
{
    memberId = 77777
    actor = 55555
    feed = article
}

FollowSources
{
    memberId = 77777
    followeeID = 55555
    followCampaign666
}
Rigorously set up, then identify whether the feature increased the success metric.

Design: How long to run experiment, on whom?

Implement: properly randomizing to ensure no bias

Analyze: Go or no-go? monitor success metric, ideally automated on company-wide platform for holistic view of impacts
Iterate. How can we revise? How can we tweak to optimize?

Reporting, monitoring, ad hoc analysis

Long term measures of engagement/success

Analysis to inform revision of design
To create a data driven organization

1. Consciously map how you use data in each phase of the product lifecycle

2. Treat data as a first class citizen

3. Create a culture which expects decisions are informed by data
Invest in foundations. 

Culture

Process

Tools
Data scientist as a partner, not a service – give context and ownership

Strong bias to actionable impactful insights, speed of iteration & feedback

- Data Foundations: governed datasets, consistent shared datasets and metrics
- Data Democratization: self serve data exploration platform
- Enable innovation: environment supports speedy ad-hoc analysis
Democratize data – self served data exploration platform

Enable people in your organization (execs, product managers, designers) to have data at their fingertips – to ask and answer questions

<table>
<thead>
<tr>
<th>Build</th>
<th>Measure</th>
<th>Track</th>
<th>Ship</th>
<th>Tweak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actionable insights lead to product feature</td>
<td>Success Metric Definition</td>
<td>Tracking Instrumentation spec</td>
<td>Experimentation Setup and Analysis</td>
<td>Post-Launch Analysis</td>
</tr>
</tbody>
</table>

**Diagram:**
- A data visualization showing trends over time, with filters for domain, dataset, and metric.
• Invest in creating the right metrics
• User-centric mindset; optimize for user value not team success

All stakeholders agree upon success metrics prior to launching the feature test

• Platform of shared tiered metrics visible to entire company
• A metrics pipeline that enables easy implementation of metrics (and not manual one-offs)
Actionable insights lead to product feature

Success Metric Definition

Tracking Instrumentation spec

Experimentation Setup and Analysis

Post-Launch Analysis

Build  Measure  Track  Ship  Tweak
Data as a first class citizen. Feature excellence for users, data excellence for employees.
- Data tracking bugs as a launch blocker

Proactive joint definition of data requirements and contract (schemas) by producers and consumers
- Data Model Review Committee

Data spec tool for source of truth
- Data Quality Monitoring tool to ensure data contract is met
- Automated testing of tracking
- Belief in proactive controlled experiment as decision making tool
- Incentives to innovate via experimentation and move key metrics

- Efficient and principled lifecycle, from inception to decision
  - Before: Review experiment design and implementation
  - After: Experiment review meeting: stakeholders discuss impacts and implications

- Company-wide platform for experimentation, with tiered key metrics
- Automated metric reporting & analysis capability; limit ad-hocs
Actionable insights lead to product feature

- Success Metric Definition
- Tracking Instrumentation spec
- Experimentation Setup and Analysis
- Post-Launch Analysis

Build
Measure
Track
Ship
Tweak

Page Views & Uniques

Groups Pageviews
-0.5%
±0.7%

Groups Uniques
-0.1%
±0.1%

Home Page Pageviews
-0.1%
±0.2%

Home Page Uniques
+0%
±0%

Groups Uniques Results

% Delta
0.1% ± 0.3%
p-value
0.70

Variant
treatment control
Average
0.098792 0.098865
Sample Size
9,995,673 9,995,897

April 7, 2014
Iteration and innovation

Metrics meeting: weekly to understand performance and product value

Long term hold-out groups for monitoring impacts
This talk will be a success if we:

1. Review the steps of data driven product innovation

2. Understand what is needed to best enable fostering of (or transformation into) a data-driven organization: culture, process, and tools.
To create a data driven organization

1. Consciously map how you use data in each phase of the product lifecycle

2. Treat data as a first class citizen

3. Create a culture which expects decisions are informed by data
Thank you.

http://www.linkedin.com/in/yaelgarten
@yaelgarten