Who are we?

Josh Poduska
Chief Data Scientist
Domino Data Labs

Patrick Harrison
Director of Data Science
S&P Global Market Intelligence
Agenda

Introduction and welcome
Motivation: why this matters
Common challenges to managing data science in the enterprise
Guiding principles and framework
Process
  Breakout Exercise: Project pre-flight checklist
Break
People
  Breakout Exercise: Team-building plan
Managing Technology and X-Factors
Summary
Data Science

Why is it different; why does this matter?
At the heart of data science lies an innocuous sounding thing...
called a model.
The implications of not becoming a Model-Driven Business are existential.
Breakthroughs open new revenue streams, expand into new markets, create and deliver new products.

Operational efficiency gains that compound through constant incremental improvement.
Jeff Bezos’s 2016 Annual Letter to Shareholders:

At Amazon, we’ve been engaged in the practical application of machine learning for many years now. Some of this work is highly visible: our autonomous Prime Air delivery drones; the Amazon Go convenience store that uses machine vision to eliminate checkout lines; and Alexa, our cloud-based AI assistant.

But much of what we do with machine learning happens beneath the surface. Machine learning drives our algorithms for demand forecasting, product search ranking, product and deals recommendations, merchandising placements, fraud detection, translations, and much more.
And Yet…

90% of companies want to make data science an operational part of their business

30% have 5+ models in production
Many organizations treat data science as a technical practice — instead of an organizational capability.
How most organizations operate

Slow, unclear path

Ad-hoc, one-off production

"black hole"

Data Science Leaders

Risk and Compliance

Business stakeholders and IT system owners

Data Scientists

CIOs and IT leaders

Business Leaders & Data Science Leaders

Wild west of desktop data science

@#!

Slow, unclear path

$$$$ — ?
How model-driven organizations operate

Model Development
- Data Scientists
- Business leaders & Data science leaders

Validation & Review
- Data science leaders
- Risk and compliance leaders
- Business stakeholders and IT system owners

Model Delivery/Deployment
- CIOs and IT leaders

Monitoring & Feedback
- Business leaders & Data science leaders
Barriers
Data-Era Infrastructure Mentality

Data science demands new degrees of infrastructure flexibility and scalability.
Garage Silos

Data scientists’ work is bespoke, ad hoc, and siloed.
Companies struggle to put models and model-backed products into production. Or if they make it into production, companies struggle to measure their impact and drive subsequent improvement.
Model Liability

Models built without proper checks and controls have the potential to do significant harm to a company’s profits, brand, and reputation.
Mindsets of the most effective data science organizations

- Reusable knowledge > Producing an answer
- Speed of iteration > Big breakthroughs
- Tool agility > Any one tool
- Process and culture > Any one piece of technology
A framework for managing data science as a capability

**Process**
- Deliver measurable, reliable, scalable outcomes

**People**
- Attract, hire, onboard, retain, and organize world-class talent

**Technology**
- Productivity and best practices to enable scale

**X-Factors**
- Managing model liability
- Navigate organizational politics
Process
Process

- Deciding what we do
- Doing projects
- Wrapping up projects
Deciding what we do: *engage the business*

**Typical approach**

- Data -> Analysis -> Product Development -> KPI
- Common Pitfalls = Scope creep, loss of stakeholder enthusiasm, no crisp measure of success

**Better method**

- Problem -> Relevant KPIs -> Product Requirements -> Analysis Necessary -> Data
- Result = Greater focus, lower risk

- Business process map
- Educate stakeholders on what is possible (avoid perception of magic)
- Allow all stakeholders to submit ideas
- Publish monthly to all stakeholders, re-prioritize at least quarterly
Project Prioritization

- Calculate Value at Stake
  - Order of magnitude value capture ($100k, $1mln, $10mln, etc.)
  - How much improvement is realistic?

- Estimate Effort
  - Order of magnitude cost estimation (1hr, 1 day, 1 week, 1 quarter, 1 year)

- Forecast Risks
  - Barriers to adoption
  - Potential consequences of errors or performance degradation

![Graph showing Do! and Don’t!]
Prioritization Pitfalls

- Embark on never-ending science projects
- Overlook linkages between model insight and business action
- Focus on what’s easy or clever instead of what’s valuable
- Cost estimates fail to consider integration, maintenance, retraining
Project kick-off

• “We don’t fail because of the math… we fail because we don’t anticipate how the math will be used.”

• Time saved here pays 10x in development and 100x in prod

• “Product management” principles apply to data science projects just as much as engineering projects
Kickoff checklist

• Business case definition
• Stakeholder mapping
• Technology needs
• Data availability
• Prior art review
• Model delivery plan
• Success measures
• Compliance and regulatory checks
## ROI Math Example

<table>
<thead>
<tr>
<th></th>
<th>PROJECT #1 – CHURN PREDICTION</th>
<th>PROJECT #2 – FRAUD CLASSIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Value at Stake</strong></td>
<td>100,000 customers *</td>
<td>50,000 applications *</td>
</tr>
<tr>
<td></td>
<td>$1000 ARR *</td>
<td>1% fraud rate *</td>
</tr>
<tr>
<td></td>
<td>10% current churn</td>
<td>$2000 avg. resolution cost</td>
</tr>
<tr>
<td></td>
<td>= $10mln problem</td>
<td>= $1mln problem</td>
</tr>
<tr>
<td><strong>Potential for Improvement</strong></td>
<td>Low (already quite low churn)</td>
<td>High (not doing anything today, no headcount )</td>
</tr>
<tr>
<td><strong>Dependencies</strong></td>
<td>Enough support staff to act?</td>
<td>App dev team integration</td>
</tr>
<tr>
<td><strong>Level of Effort</strong></td>
<td>1 month</td>
<td>1 quarter</td>
</tr>
<tr>
<td><strong>Risk of False Positive</strong></td>
<td>Low (extra support)</td>
<td>High (bad customer experience)</td>
</tr>
<tr>
<td><strong>Risk of False Negative</strong></td>
<td>Medium (lost revenue)</td>
<td>High (more lost revenue)</td>
</tr>
<tr>
<td><strong>Re-training Requirements</strong></td>
<td>Medium (marketing mix changes slowly)</td>
<td>High (adversarial)</td>
</tr>
<tr>
<td><strong>Change Management Requirements</strong></td>
<td>Low (educate support team, currently use random / intuition)</td>
<td>High (modify real-time application flow)</td>
</tr>
</tbody>
</table>
Stakeholder Mapping

• Best Practices
  • Define responsible parties from each group: data science, business, DevOps, application dev, compliance, etc.

• Common Pitfalls
  • Lack empathy with goal of actual end user
  • Throw results “over the fence” to IT with no context
Technology Needs

• Best Practices
  • Consider opportunities to accelerate research
  • Identify dependencies early

• Common Pitfalls
  • “One size fits all” tooling
  • Underpowered infrastructure
Data Availability

• Best Practices
  • Leverage existing sources first to build baseline
  • Create synthetic data with realistic characteristics
  • Track engagement with datasets to automatically discover experts

• Common Pitfalls
  • Wait for “perfect” data
  • Buy external data without clear onboarding plan
Prior Art Review

• Best Practices
  • Review state of the art — internally and externally

• Common Pitfalls
  • Culture of NIH
  • Nose-to-the-ground mindsets
  • No single source of truth
Model Delivery Plan

• Best Practices
  • Design multiple mock-ups of different form factors
  • Designate approvers in advance (IT, DS, biz)
  • Create process flow to precisely show where model will impact
  • Consider agile approach

• Common Pitfalls
  • Fail to educate end-users
  • Over-engineer relative to the requirements use case
Success Measures

• Best Practices
  • Pre-emptively answer “how will we know if this worked?”
  • Frame in terms of business KPIs not statistical measures
  • Define needs for holdout groups, A/B testing, etc.

• Common Pitfalls
  • Not knowing when it is “good enough”
  • Fail to establish testing infrastructure and culture
Risk Mitigation

• Best Practices
  • Consider consequences of errors (e.g., false positives / negatives)
  • State likely biases in training data
  • Track ongoing usage to prevent inappropriate consumers

• Common Pitfalls
  • Assume no regulation today will last
  • Conflate model interpretability with model provenance
  • Model misuse
Wrapping up projects

• Best Practices
  • Defend the scientific method
  • Store positive and negative results
  • Preserve synthesis, intermediate results, code, data, and environment

• Common Pitfalls
  • Repeated quiet failures
  • Old analysis doesn’t run
Group Exercise #1: Fill out a pre-flight checklist for one of your projects

- Spend 15 minutes filling out template
- Discuss in groups of five for 20 minutes
# 8 Factor Pre-Launch Checklist: Questions to Ask

| Business Case | What's the desired outcome of the project, in terms of a business metric?  
What are the linkages from your project to impacting that ultimate business metric?  
What is the order-of-magnitude ROI and cost? |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Stakeholders</td>
<td>Lead DS, Proposed Validator, Data Engineers, Product Manager, Business Executive, Business End User (internal or external), Application Developer, DevOps Engineer, Compliance</td>
</tr>
</tbody>
</table>
| Technology      | What compute hardware and software infrastructure do you anticipate being necessary?  
Would your project benefit from specialized or parallelized computing? |
| Data            | What relevant data exists today? Who are the subject matter experts?  
What other data would we potentially want to capture, create, or buy externally? |
| Prior Art       | Who has worked on this business topic before (internal and external)?  
Who are the relevant experts in the techniques I will likely use? |
| Model Delivery  | Who will consume this and what form factor will the final product take (report, app, API)?  
What dependencies or resources will you require to deliver work this way (e.g., IT)?  
What are other possible delivery mechanisms, especially ones that are lighter weight or easier to test first?  
What user training is necessary to ensure adoption? |
| Success Measure | How will you know if it's working as expected, or otherwise get feedback?  
What's your “monitoring” plan, even if it's manual and subjective? |
| Risk Mitigation | How could this model be mis-used by end users?  
Any constraints on modeling approach (e.g., interpretability requirements)? |
A note for early teams

- Don’t be overwhelmed into paralysis by complex process
- Look for low-hanging fruit to buy political capital for more headcount and risky projects
- Find senior sponsor
- Most important takeaway: engage the business as partners early and often
Process: The Data Science Lifecycle
Common pitfalls in the data science lifecycle
### Bottlenecks and pitfalls have quantified negative impact

<table>
<thead>
<tr>
<th>CHALLENGES</th>
<th>DATA SCIENCE BOTTLENECKS</th>
<th>BUSINESS IMPACT</th>
</tr>
</thead>
</table>
| **1** Inconsistent Project Priority and Kickoff | • Duplication of work wastes time and slows down progress  
• Inability to leverage past work and customize across locations  
• Scope creep and loss of stakeholder enthusiasm                                                                                                                                                                                                                                          | • X% of projects have little/no impact  
• Y number of weeks lost by employees identifying what projects have been done before and understanding that work                                                                                                                                                                 |
| **2** No Access to Technology On-Demand | • <Subsidiary> approval for required infrastructure takes weeks per project  
• Insufficient infrastructure prevents differentiated innovation                                                                                                                                                                                                                               | • 4-6 weeks delays for resource requests spread between approvals and implementation  
• X time wasted replacing Data Scientists                                                                                                                                                                                                                                                         |
| **3** No Ability to Easily Deploy Results to Business | • Data Scientists waste time on mundane tasks to expose models  
• Business stakeholders complain about lengthy delays to business value                                                                                                                                                                                                                               | • Z time lost by employees setting up dashboard servers                                                                                                                                                                                                                                               |
| **4** Failure to preserve knowledge upon completion | • Lack of documentation and reproducibility of code hurts iteration  
• Projects just fade away, so null results aren’t known for future collaborators                                                                                                                                                                                                                                                                   | • Model iteration velocity slowed by average of 1m                                                                                                                                                                                                                                                     |
People
Why focus on people?

• Talent gap commonly cited as obstacle to being model-driven

• Typical tenure <2 years with 3+ month ramp

• Overwhelmed by resumes, underwhelmed by output
Framework for People

• Attract – How to lure the best talent
• Assess – Hire systematically
• Train – Focus on mindset, not just skills
• Retain – Build community and mentorship
• Organize - Define optimal roles and structure
Attracting the best and brightest

• Best Practices
  • Have a differentiated offering and strategy
  • Advertise projects, not just the company
  • Offer modern tools and commitment to open source

• Common Pitfalls
  • Write unrealistic job descriptions
  • Seek PhDs when need hackers (or vice versa)
Assessment

• Best Practices
  • Be systematic: identify required attributes, design assessments for each
  • Be analytical: track interviewer and interview type efficacy
  • Include EQ and non-technical assessments
  • Sell while assessing: simulate real work
• Common Pitfalls
  • Over-rely on tech screens
Training

• Best Practices
  • Reinforce mindsets, not just skills
    • Develop culture of reuse, compounding
  • Reward community-enhancing behavior
  • Provide “soft” skills training
• Common Pitfalls
  • “Not built here” mentality
Set expectations on time allocation upfront

- **Best Practices**
  - Emphasize listening to stakeholders
  - Compensate team on new and existing work, not just current projects
- **Common Pitfalls**
  - Employee churn from flawed expectations

Source: Max Shron, Warby Parker
Metrics of managing data science

• Best Practices
  • Share accountability with the business’s KPIs
  • Focus on iteration velocity
  • Systematically capture stakeholder feedback and engagement

• Common Pitfalls
  • Measure everyone but yourself
  • Over-index on any one project vs. factory performance
### The many hats of data science

<table>
<thead>
<tr>
<th>ROLE</th>
<th>PRIORITIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scientist</td>
<td>Generating and communicating insights, understanding the strengths and risks</td>
</tr>
<tr>
<td>Data Infrastructure Engineer</td>
<td>Building scalable pipelines and infrastructure that make it possible to do the higher levels of needs.</td>
</tr>
<tr>
<td>Data Product Manager</td>
<td>Articulating the business problem, translating to day-to-day work, ensuring ongoing engagement.</td>
</tr>
<tr>
<td>Business Stakeholder</td>
<td>Vetting the prioritization and ROI, providing ongoing feedback</td>
</tr>
<tr>
<td>Data Storyteller</td>
<td>Creating engaging visual and narrative journeys for analytical solutions</td>
</tr>
</tbody>
</table>
Organizational Design Dilemmas

• Best Practices
  • Solve prioritization and delivery problems first
  • Bridge silos with cross-cutting platforms

• Common Pitfalls
  • Fail to evolve structure as org matures
  • Confine teams to ivory tower innovation labs

<table>
<thead>
<tr>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CENTRALIZATION</strong></td>
<td><strong>DECENTRALIZATION</strong></td>
</tr>
<tr>
<td>• Community and mentorship</td>
<td>• Stronger alignment with business processes and priorities</td>
</tr>
<tr>
<td>• easier transparency for managers and IT</td>
<td>• Easier change management</td>
</tr>
<tr>
<td>• More passive technical knowledge sharing</td>
<td>• Less technical knowledge compounding</td>
</tr>
<tr>
<td>• Isolation on data science island</td>
<td>• Harder to codify best practices</td>
</tr>
<tr>
<td>• Loss of credibility with business</td>
<td>• Risk of IT governance issues</td>
</tr>
<tr>
<td>• Frustrated data scientists</td>
<td></td>
</tr>
</tbody>
</table>
What Org Design is Right For You?

- Prioritize stakeholder proximity early if internal use cases
- Tie to engineering if primary building model-driven external-facing products
- Develop hub-and-spoke as you scale
## Hiring and Ramping Plan Template: Questions to Ask

### Attracting Talent
- What’s your differentiated value proposition for candidate data scientists? List three things that make the opportunity unique, that you think will resonate with your target candidate pool.
- What are 1-3 risks that might make the opportunity less appealing than competitive opportunities? How can you mitigate or get ahead of them?

### Hiring Process
- What are the three most important attributes for your candidates? What is your assessment plan for each?

### Onboarding
- What outcomes need to have been achieved in the first 30, 60, and 90 days?
- What are the most important pieces of “tribal knowledge” your new hire needs to know, and how will she learn them? Examples include data sources, project methodologies, stakeholder dynamics, notable wins / losses, etc.

### Retention and Management
- What skills do you hope this candidate develops over the first year?
- What metrics will determine success of this candidate after a year? Examples include certain business metrics, community contributions, number of insights produced, or project iteration velocity.
Group Exercise #2: Build your hiring and ramp plan

• Spend 15 minutes filling out template
• Discuss in groups of five for 20 minutes
Technology

Collaboration
- Shared context
- Discussion
- Knowledge Management (search & discovery)

Reproducibility & Reusability
- Code
- Data
- Results
- Environments

Agility & Iteration
- Experimental agility
  - Tools / packages
  - Compute
- Deployment agility
  - Expose work back to the business quickly
Strategy: incentivize best practices “bottom up”

DATA SCIENTISTS: “I’M MORE PRODUCTIVE!”

LEADERS: “CENTRALIZED WORK!”

Test Ideas Faster

Powerful Collaboration Features

Deploy and Share Work Easily

Version Control & Reproducibility
How we approached this

Workbench

Entice data scientists with:
- Vertical and horizontally scalable infrastructure
- DevOps automation
- Computational lab notebook to track results

Collaboration Hub

Centralizing work makes it possible to:
- Find, reuse, reproduce, and discuss work.

Publishing & Deployment

Decrease time to business impact:
- Deploy models as APIs
- Deploy apps (e.g., Shiny) & reports to non-technical stakeholders
- Scheduled jobs for ETL, reporting, model retraining
X-Factors
Model liability

• Problem emerges at later maturity
• Track and guardrail model usage
• Document risks and trade-offs made in flight, not post hoc
• Pre-emptively establish validation, monitoring, and compliance controls
Navigating organizational politics

• Educate executives on reality of probabilistic research
• Anticipate demands of procurement (ROI of aggregate project portfolio)
• Frame impacts of data science investment:
  • Out-compete peers
  • Increase operational efficiency
  • Reduce costs (headcount etc)
  • Reduce risk
Summary
Summary

• Data science success is not adding up individual successes, it’s an *organizational capability*
• Alignment and partnership with the business is critical
• Process – Enforce a pre-flight checklist
• People – Develop hiring and onboarding plans
• Technology - Leverage technology to increase productivity and best practice processes
• X-Factors – Navigate politics and risk
Struggling with your own lifecycle?

• Ask us about Domino’s Data Science Lifecycle and Value Assessment offerings
  • Tailored analysis of existing processes, gaps, and tangible best practices
  • Leverage our ROI analysis templates across your portfolio
Want to learn more?

• Check out this content for more information
  • The Practical Guide to Managing Data Science at Scale
  • Data Science Management Survey Report

• Stop by our booth #1403

Questions?