



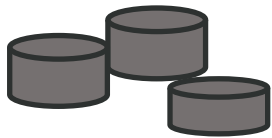
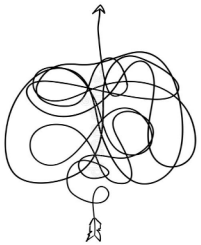
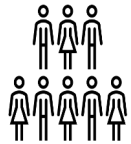
Solve the Analytics Last-Mile
Problem with a DataOps
Process Hub

DataKitchen DataOps Process Hub Agenda

- Current Situation of Business Analytics (BA) Teams
- DataKitchen DataOps Process Hub for BA Teams
- Examples of BA Teams Using DataKitchen's DataOps Process Hub
- DataKitchen DataOps Process Hub Demo Example
- Conclusion & Questions

Business Analytics: The Problem

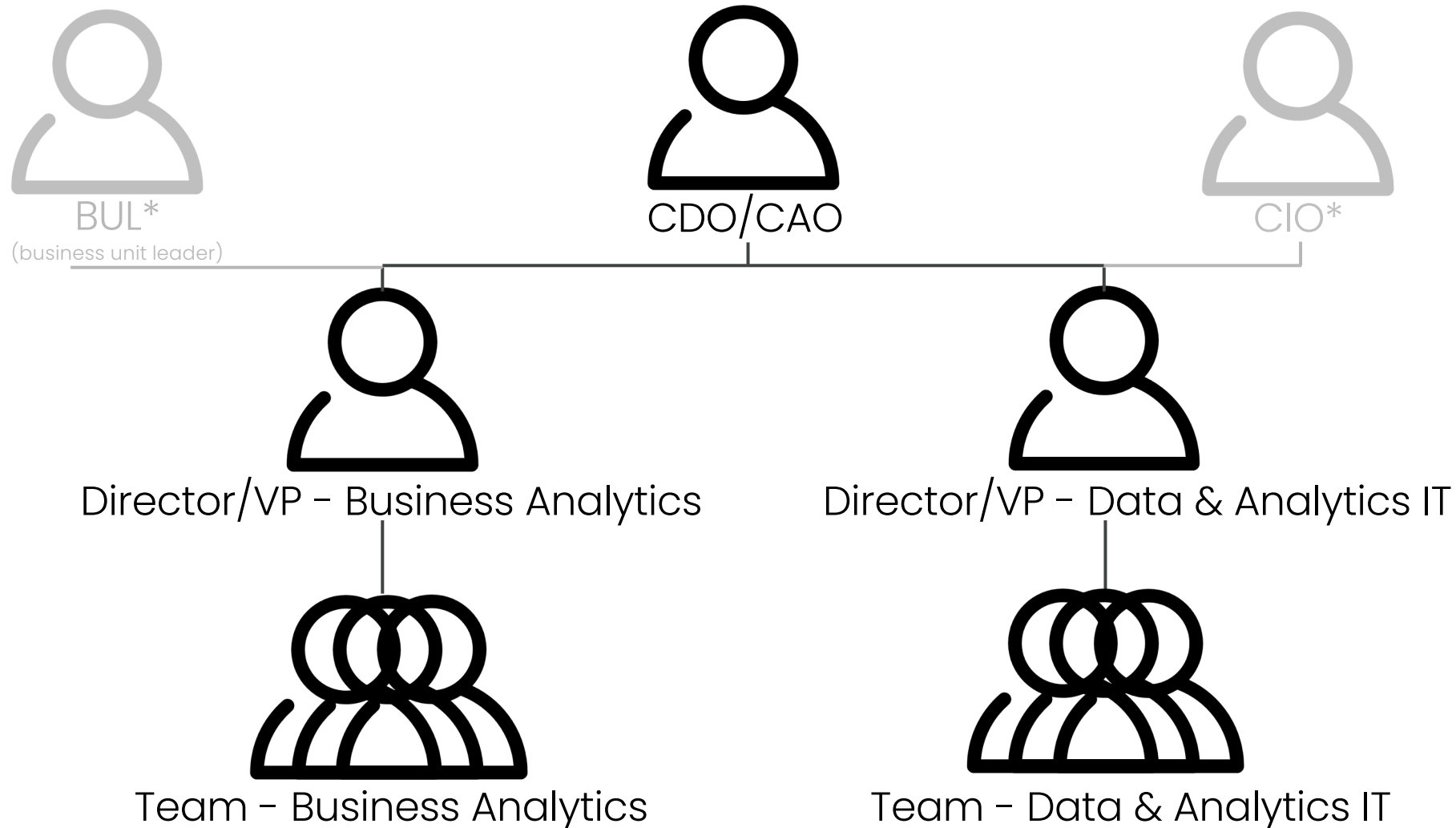
Business
Analytics
Analysts



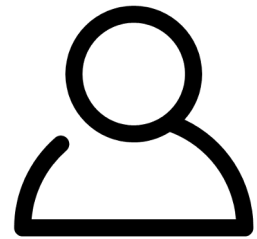
Base Data
(Data
Lake/Warehouse)

- Business Analytic Teams **MUST** deliver value to their customer, make rapid changes, and have ongoing, updated production deliverables
- Lots of data tables from IT and diverse data cause Business Analytics (BA) analysts to do a **huge amount of fragile, error prone, and high effort data work** to get insight
- Repetitive data processes and query work **reduces the time the Business Analytics team needs create original insight**
- Inconsistencies in how the analytics team applies business data process and business logic cause errors. **Business Analytics firefighting and stress are common**
- Lack of a tailor-made data representation, up to date documentation, and reusable business logic cause analysts extra time to do ad hoc analysis and make training new Business Analytics analysts difficult.
- New insight requires manual time. Reuse is low and costs are high.
- So, you can either continuously throw BA staff at the problem or work with IT and take many months to build something that is costly to change and out of date on arrival

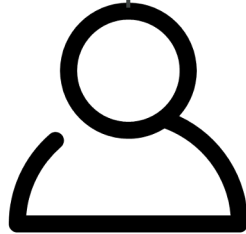
Sometimes Data & Analytics Teams Are Together



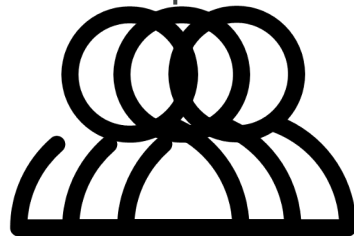
Sometimes Data & Analytics Teams are Apart



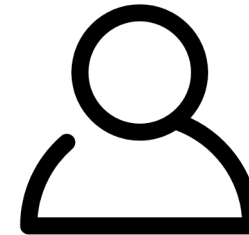
BUL*
(business unit leader)



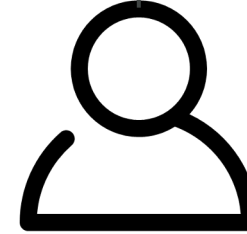
Director/VP - Business Analytics



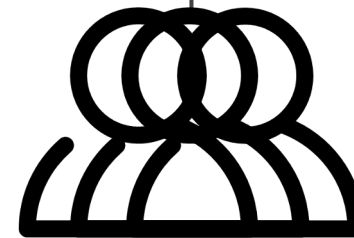
Team - Business Analytics



CIO*

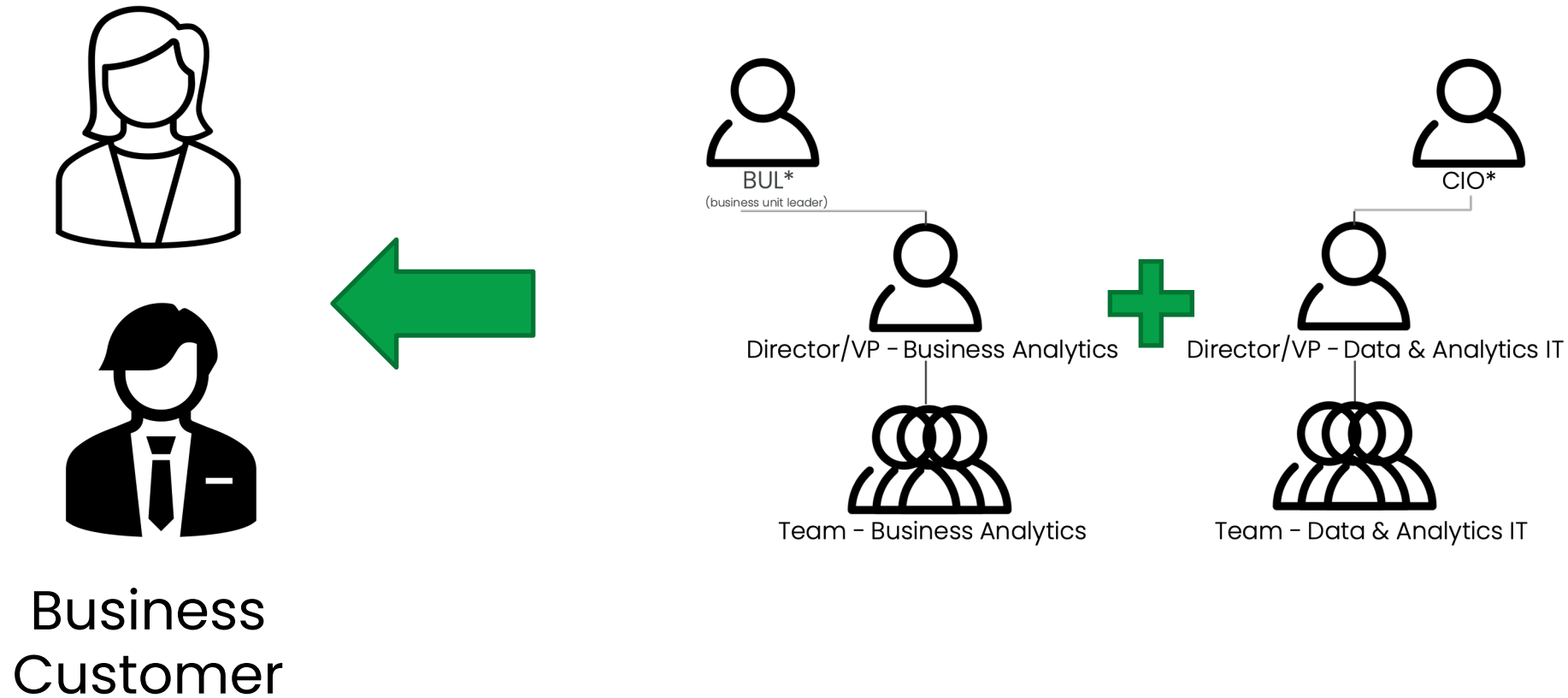


Director/VP - Data & Analytics IT

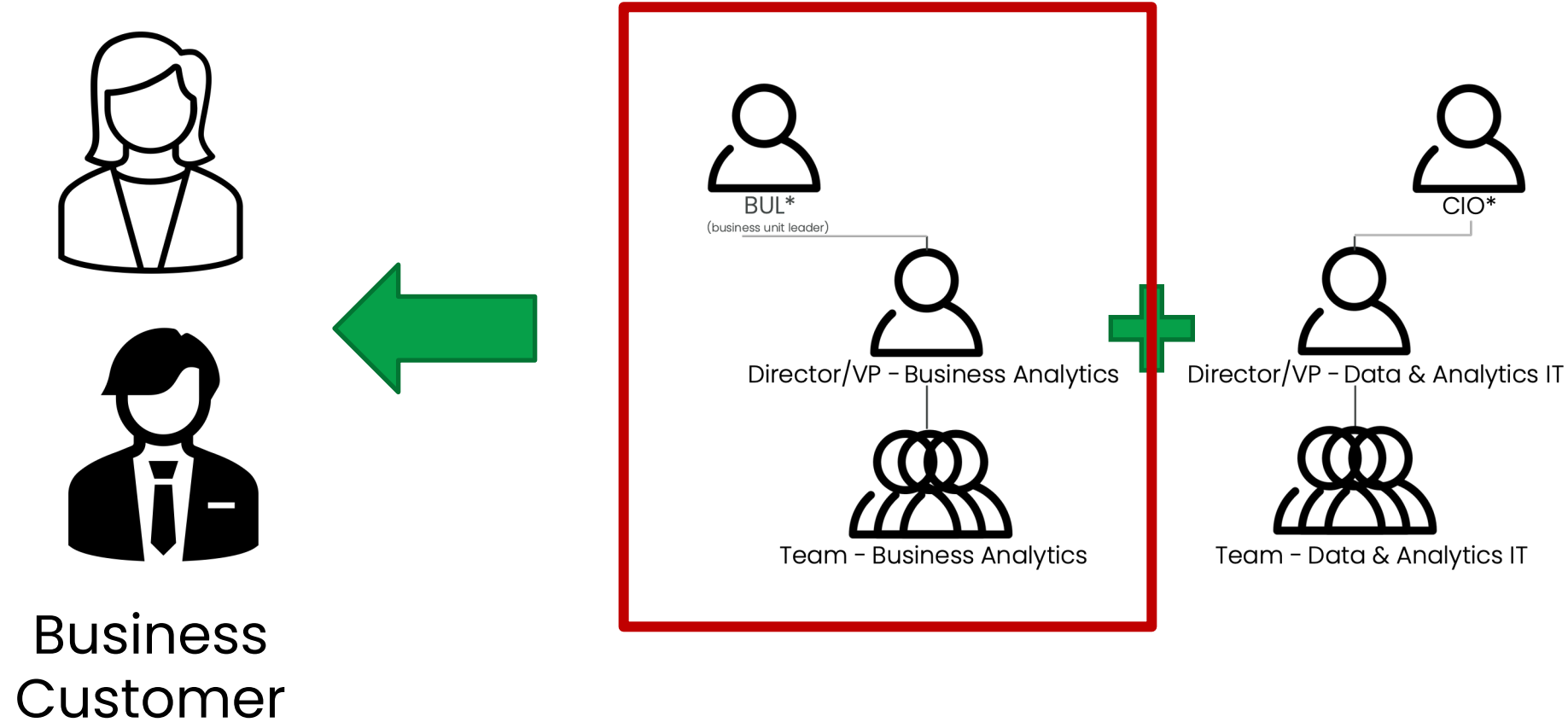


Team - Data & Analytics IT

Teams Apart Deliver Joint Value



Unique Problems of Business Analytic Teams



Business Analytic Team IT Challenge

Some Data & Analytic IT teams are **great!**

- They want the Business Analytic (BA) Team to be successful
- They are focused on BA Team success
- They work in an Agile & iterative way
- They judge their success by the success of the BA Team

However, some Data & Analytic IT teams are **not so great**

- They 'throw it over' the fence to the BA Team
- They are not focused on delivering customer value but on their own internal process
- They are 'value avoidant'
- They take many, many months to get anything done
- When they do get something done it takes lots of re-work to get it right
- BA may not trust the result & have to do their own QA
- BA still needs the IT team to provide access to the data!

DataKitchen DataOps Process Hub Agenda

- Current Situation of Business Analytics (BA) Teams
- DataKitchen DataOps Process Hub for BA Teams
- Examples of BA Teams Using DataKitchen's DataOps Process Hub
- DataKitchen DataOps Process Hub Demo Example
- Conclusion & Questions

Typical Business Analytics Quotations

"How do we do the same level of service for our business customers, faster?"

*"I wish I were in a better position.
I got cut off at the knees from a data perspective,
& I am getting handed a sandwich of sorts.
And not a good one!"*

"What kind of back-end process can we have to answer what-if questions faster?"

"My team is being too reactive. Need a system to help them be more proactive with our customers."

"The design of Data Warehouse was built to ease IT data loading & not Business Analyst insight."

"This is a last-mile problem."

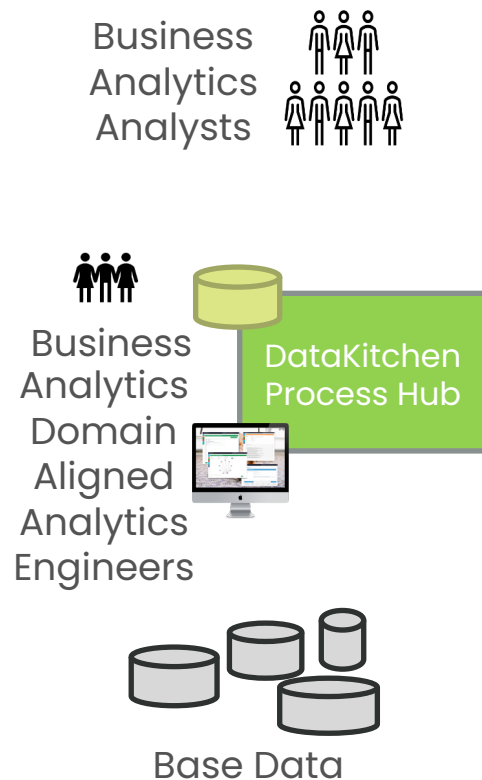
"Right now, delivering insight to the business is a 90% Business Analytics problem & 10% IT problem."

*"Hard to change the Data Warehouse.
I don't understand how they do it or QA."*

"We spend 80% getting data together, 20% on insight."

*"Can we create data in an 'answer-ready' format?
That has data quality & can answer many scenarios?
Where we can just click a few buttons?"*

Solution: The DataKitchen DataOps Process Hub

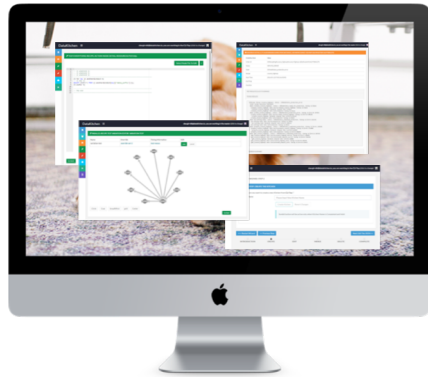


Solution: Process Hub

- Instead of 1000 tables give each Business Analytics (BA) Team **a dozen tailored tables they need to get trusted insight.**
 - Include clear, up-to-date dictionaries so that new BA Analysts can get up to speed rapidly.
 - In place of a multi-month development cycle, give each team the control of those tables & the ability to **make rapid daily/weekly changes** to the data & schema in those tables with high quality.
- Give each team a **small number of aligned Analytics & DataOps Engineers** who learn the data, the business problem & can be a partner with the BA Analysts in making them successful.
 - Automate the hell out of the process so that the Data Engineers spend no time maintaining the answers to the last question & have all their time devoted to answering the next Business Analytic question.
- Use the Process Hub to store all business logic in GitHub for reuse, sharing, best practices & staffing flexibility.
 - Store Business Analytic intellectual property in the Process Hub (i.e., Domino models, SQL, reports, tests, scripts, etc.).

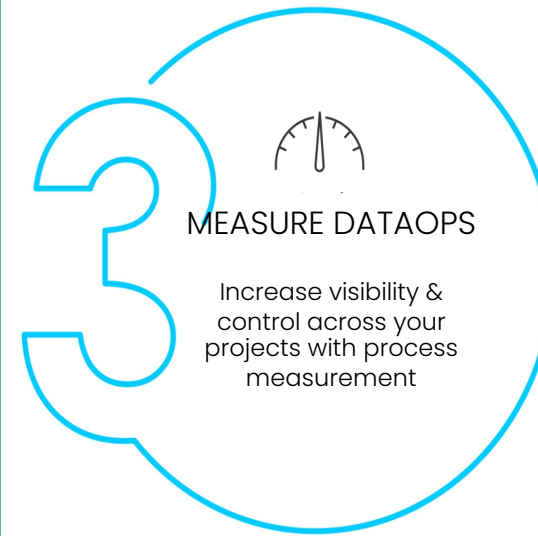
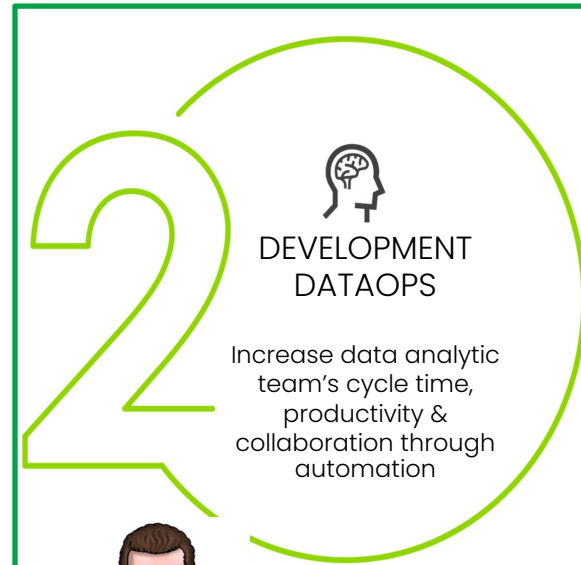
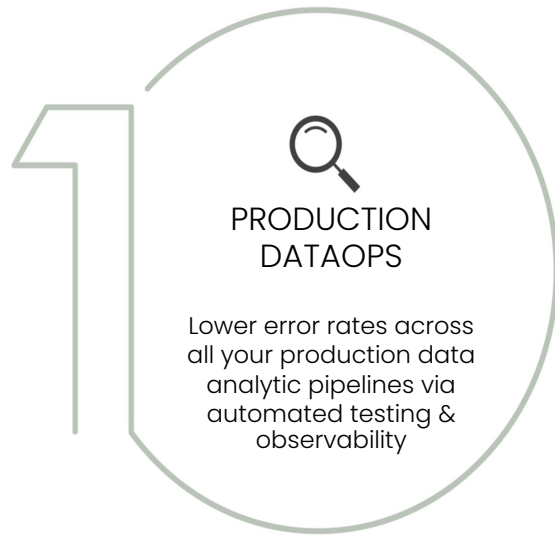
DataKitchen DataOps Process Hub: Benefits

FEATURE	BENEFIT
Data designed for Business Analytics	• Significantly faster time-to-insight for ad hoc requests
Automating (& reusing) processes with a button push	• Significantly fewer dollars spent on technical mechanics & re-work • Reallocate Business Analytics consultants to more valuable tasks
Enablement for Business Analytics, NOT IT infrastructure	• Extend Data Lake/Warehouse investment, not replace



Replace an army of people with DataKitchen's DataOps Process Hub to allow Business Analytics Teams to **produce more insight, faster, at the speed your customers & business demand**

DataKitchen DataOps Process Hub: Team



Chris
DataOps
Engineer



Priya
Data Scientist
Business
Analyst



Ahmed
Analytics
Engineer

Chris: Operations Optimizer
Goals: Set up & maintain environments; speed & ease of deployment; lower cycle time; & make data teams awesome!

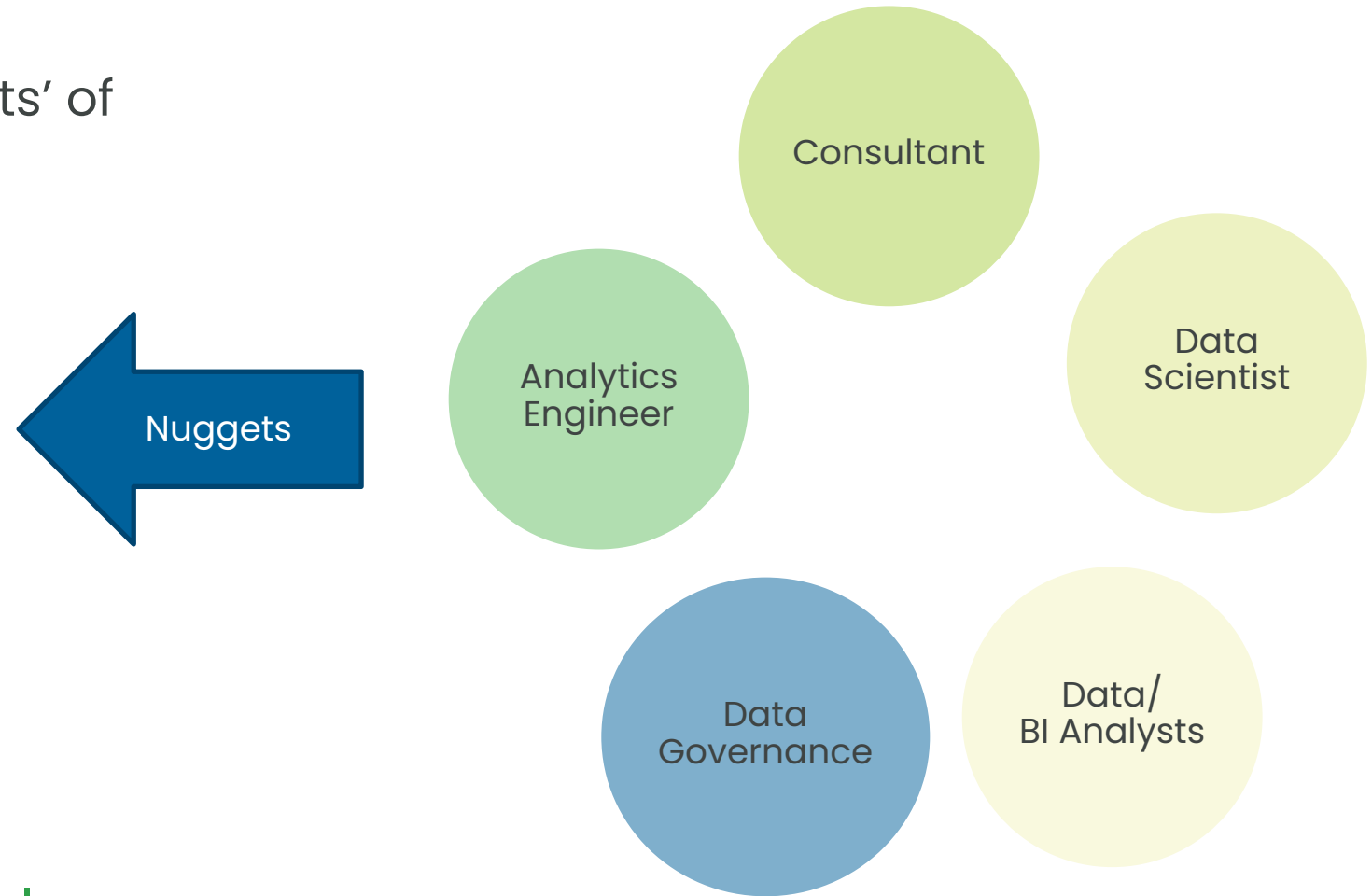
Ahmed: Data Doer
Goals: Build Data sets that are ready for insight generation & minimize data errors.

Priya: Viz. Victor (or Data Scientist, Viz Specialist, BI/Analyst, etc.)
Goals: Create cool insight for customers & innovate with data

DataOps Engineering Is about Collaboration through Shared Abstraction

DataOps Engineers take 'nuggets' of code (i.e., ETL, SQL, python, xml, tableau workbooks, etc.) &:

- Put into pipelines
- Create tests
- Run the factory
- Automate deploys
- Work across people
- Measure success
- Enable DataOps self services

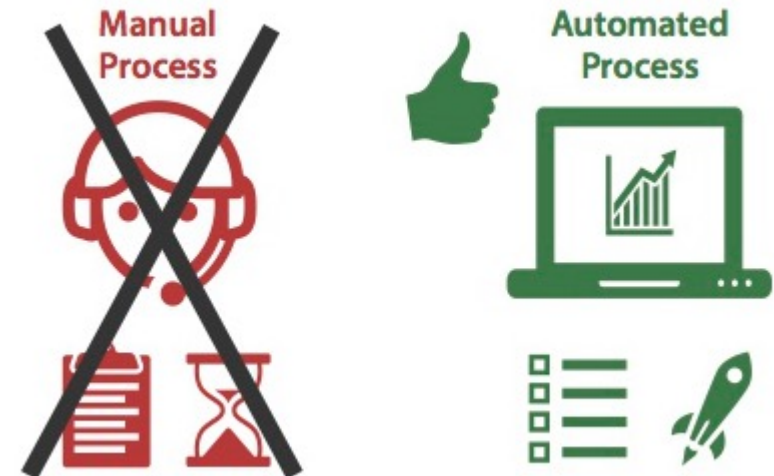


Make what is invisible, visible.

DataOps Engineering Is about Automation

Automate These Tasks:

1. Production orchestration
2. Production data monitoring / testing
3. Self-service environments
4. Development regression & functional tests
5. Test data automation
6. Deployment automation
7. Shared components
8. Process measurement



Analytics (Data) Engineer in a Process Hub

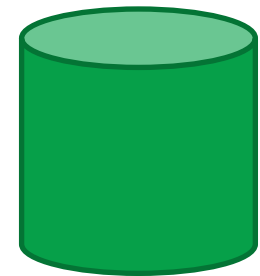
Make the next urgent deadline easier to meet



Ahmed
Analytics
Engineer

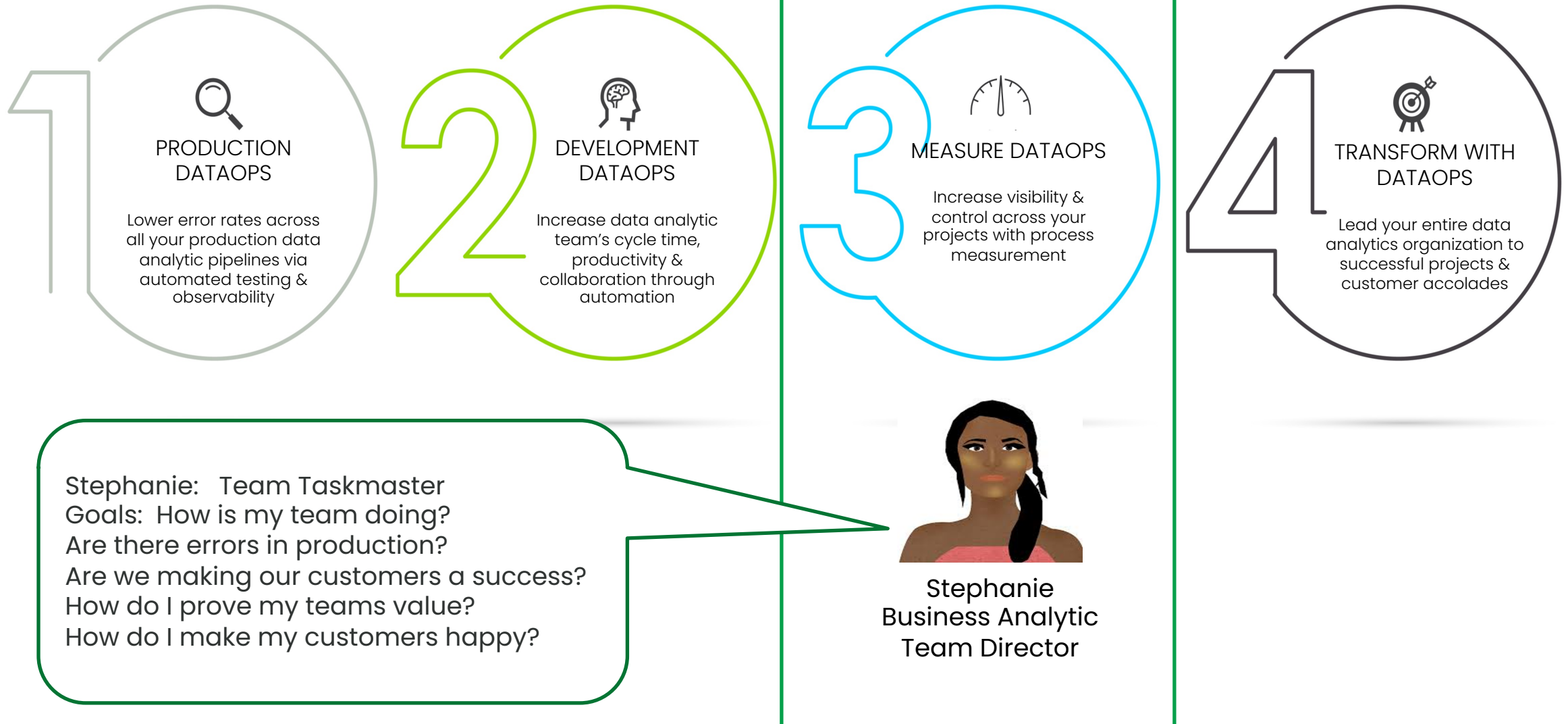
Role of the Analytics Engineer with a DataOps focus:

- Build useful, ready-to-analyze data sets
- Don't just run to the next fire
 - Ease the path to production & ensure scalability
 - Streamline processes for automation, efficiency & reuse
 - Move Assumptions, Business Rules & Code upstream
 - More consistent, more reliable & more efficient
 - Build out a flexible framework
 - Curate & manage production processes
 - Build out a production toolkit
 - Shortcuts for future tasks
 - Collaborate over the process



SQL
ETL
ELT
Data

DataKitchen Process Hub: Team



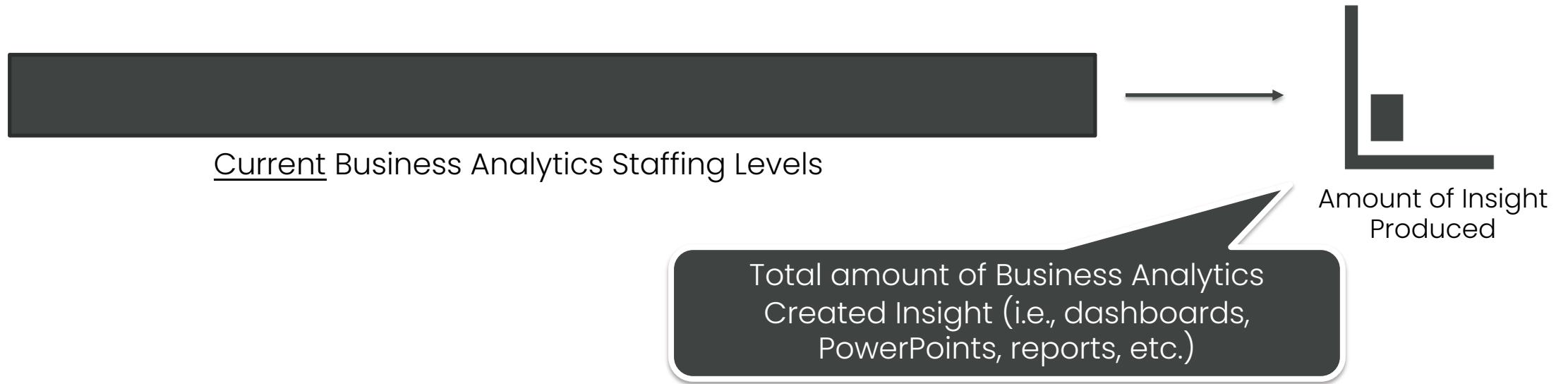
Multiple Projects, Multiple People



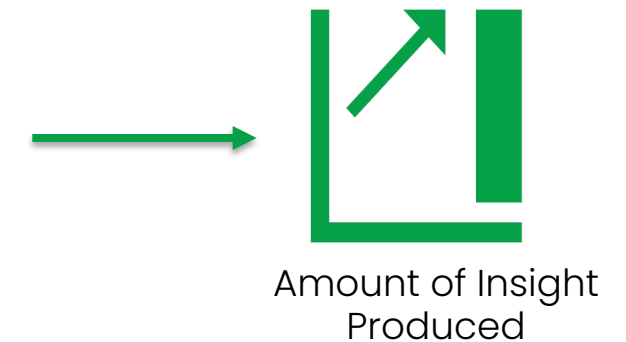
Stephanie
Business Analytic
Team Director

- Challenges:
 - Lots of projects starting & stopping
 - People moving from one task to another
 - Lots of people & vendors rotating in & out
- How do you:
 - Get consistency between projects?
 - Support collaboration?
 - Get a longer-term view?
 - Curate a series of 'shortcuts' or reusable components that can span projects & team rotation to enable efficiency?
- The Solution:
 - Everyone is focused on 'Data Hubs' to bring all the data together in place
 - You also need a Process Hub to bring all the people & processes that act upon data together in one place
 - Managing & curating processes is just as important as the data itself in bringing value to business customers

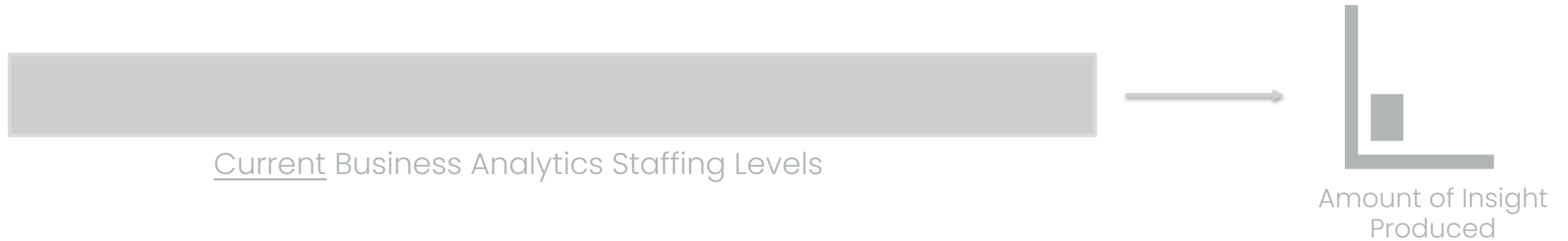
Process Hub Value – Key Question



KEY QUESTION: How to increase the total amount of Business Analytics insight generated without continually adding more staff (& cost)?



Answer: Process Hub Increases Productivity



DataKitchen DataOps Process Hub Agenda

- Current Situation of Business Analytics (BA) Teams
- DataKitchen DataOps Process Hub for BA Teams
- Examples of BA Teams Using DataKitchen's DataOps Process Hub
- DataKitchen DataOps Process Hub Demo Example
- Conclusion & Questions

Process Hubs in the Wild

What's Really Different?

Data Hub:

- A repository, exchange, collection point for data in the enterprise
 - One-way street: from raw to processed data layers

Process Hub:

- A hub for the processes that act on data
 - Works with the Data Hub
 - Data flows between teams, functions – more complicated than levels
 - Data keeps changing: valid data is a product of an effective process
 - No Single-Version-of-Truth without a Single-Version-of-Process

Three Ways a Process Hub Increases Productivity

DataKitchen's DataOps Process Hub software plus staffing Analytic Engineering makes the Business Analytics staff much more efficient in three ways:



'The Wedge': Trusted analyst-ready data sets, under Business Analytics control, with rapid changes, understandability, enables ease of answering questions with high reuse.

Benefits: lowers the cost-per-question & Business Analytics productivity rises



'The Hammer': Full automation, from source to value delivery, enabling fully zero touch end-to-end ongoing insight delivery.

Benefits: reduces recurring work & Business Analytics get more insight generating staff

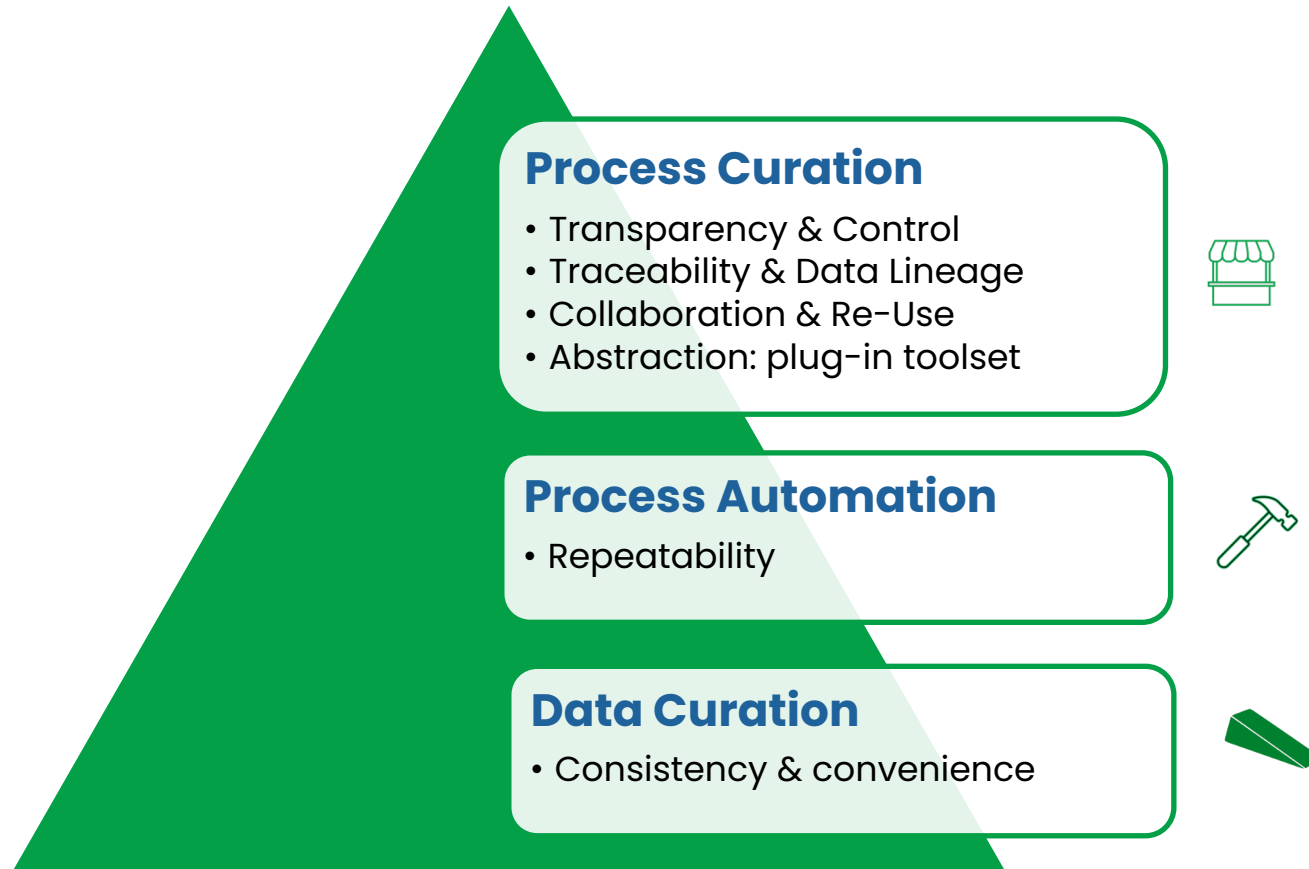


'The Store': A centralized repository under version control for reuse, sharing, best practices & storing all Business Analytics intellectual property.

Benefits: lowers risks & redundancies, & Business Analytics productivity rises

Hierarchy of Value

Increasing benefits



Process Hubs in the Wild: Its Impact on Teams

Why You Can't Leave a Process Hub to IT

- Requires business expertise
- Requires analytics expertise
- Subject to rapid change
 - Always changing conditions
 - Always new questions
- Requires meta-understanding of a process
 - Not just how to do it
 - What aspects are generalizable to what levels?

Process Hubs in the Wild: Its Impact on Teams

Why You Can't Leave a Process Hub to Analysts

- Requires enterprise priorities
- Requires cross-functional cooperation
 - Multi-team, multi-vendor
- Process-orientation vs deliverable-orientation
- Watch out for heroes & brilliant people

Process Hubs in the Wild

The Vacuum Today

Business Analyst Focus Today

- **Brand-Focused:** good but can be siloed.
- **Deliverable-Driven:** good but can be short-term.
- **Oversight-Resistant:** vendors & BA prefer less transparency, second-guessing.
- **Hero Focus:** Less incentive to build reusable, reliable enterprise features. Let's get it done.
- Relentless need to productionize technical debt



The Danger: Hidden Code, Costs, Risks of Tribal Knowledge

```
--All with new codes
WITH all_ AS
(SELECT
  _id, min(service_date) service_date
FROM (SELECT
  _id, service_date
FROM _current.raw_ _fact a
LEFT JOIN d_current.raw_ _code b
ON a._id = b._id
WHERE upper(b._code) IN ('.00', '.01', '.02', '.8', '.80', '.81',
'.82', '.9', '.90', '.91', '.92', '.00',
'.01', '.02', '.50', '.51', '.52', '.60',
'.61', '.62', '.A0', '.A1', '.A2', '0',
'1', '2')
UNION
SELECT
  _id, service_date
FROM _weekly_current.raw_ _fact
WHERE upper(_code) IN ('.0', '.8', '.80', '.81',
'.82', '.9', '.90', '.91', '.92', '.00',
'.01', '.02', '.50', '.51', '.52', '.60',
'.61', '.62', '.A0', '.A1', '.A2', '0',
'1', '2')
)
GROUP BY 1 ),
```

Code from a vendor's laptop:

- What is it?
- Where is it used?
- Who wrote it?
- What is its history?
- Is it tested?
- How do I share it?
- How do I find it?
- Can anyone see it & change it?
- Will it be lost?
- Could I just use the resulting data?
- Are these codes up to date?

Business Analytics Tribal Knowledge = high costs & slow speed to insight

Process Hub Thinking

How to Think Like an Analytics Engineer

- Break out the “recipe” steps
 - Required data sources, team touches, sequence of updates & transforms, functional tasks
- Understand the dimensions of change
 - What changes always, sometimes, never?
 - What’s structure vs. data? (i.e., constants vs variables, columns vs rows)
 - What’s a parameter in a consistent process?
- Re-sequence the critical path
 - Evolved processes, work-on-deadline: not always done in the right order
 - Move upstream: redundant tasks, assumptions, business rules
 - Keep downstream: unique requirements, specialized logic, changing variables
- Goal: balance **flexibility** & **consistency**
 - Give Analysts the levers they need to control the machine

Process Hub Thinking

What to Look for: Laptop SQL & Cubby-Hole Logic

- Routine code to transition to automation
- Manual steps to automate
- Redundant code to simplify
- Redundant tasks to consolidate
 - Similar needs from different teams
 - Point teams to existing data features & shortcuts
 - Create & evolve new features & processes
- Hard-coded logic to make data-driven
- Assumptions, business rules & code to move upstream
 - Break down team silos, even within the same vendor
- QC steps to add or automate

```
with exclusion AS
(SELECT
  _id, min(service_date) service_date
FROM (SELECT
  _id, service_date
FROM a_current.raw_
LEFT JOIN d_current.raw_
ON a._id = b._id
WHERE upper(b._code) IN
(' .10', ' .11', ' .22', ' .12', ' .21', ' .20', ' .12',
' .11', ' .1', ' .76', ' .4', ' .79', ' .83', ' .81',
' .89', ' .45', ' .4', ' .3', ' .71'))
```

```
WHERE
(
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
OR
  _name ILIKE '% %'
```



 =VLOOKUP(H3,B4:E13,2,FALSE)

Process Hub in Action

Example: Target List Management for Pharma

Goal: Identify, target & track doctors to market drugs & treatments

- Common task within a company – multiple brands
- Interplay of enterprise, division & team requirements
- Interplay of analytic & operational involvement
- Mission-critical tasks with deadline pressure

Process Hub in Action

Example: Target List Management for Pharma

Familiar Challenges

- Analytic processes must be flexible
 - Metrics, segmentation, selection
 - Predictive modeling
 - Experimentation, iterations, rapid strategic change
 - This isn't static data! This isn't IT infrastructure!
- Mission-critical pipeline must be consistent, accurate
 - Datasets are always changing: MDM, merge/purge, demographics
 - Regulatory requirements: rules for suppressions, opt-outs
 - High cost of failure
 - Potential QC nightmare

Process Hub in Action

Example: Target List Management for Pharma

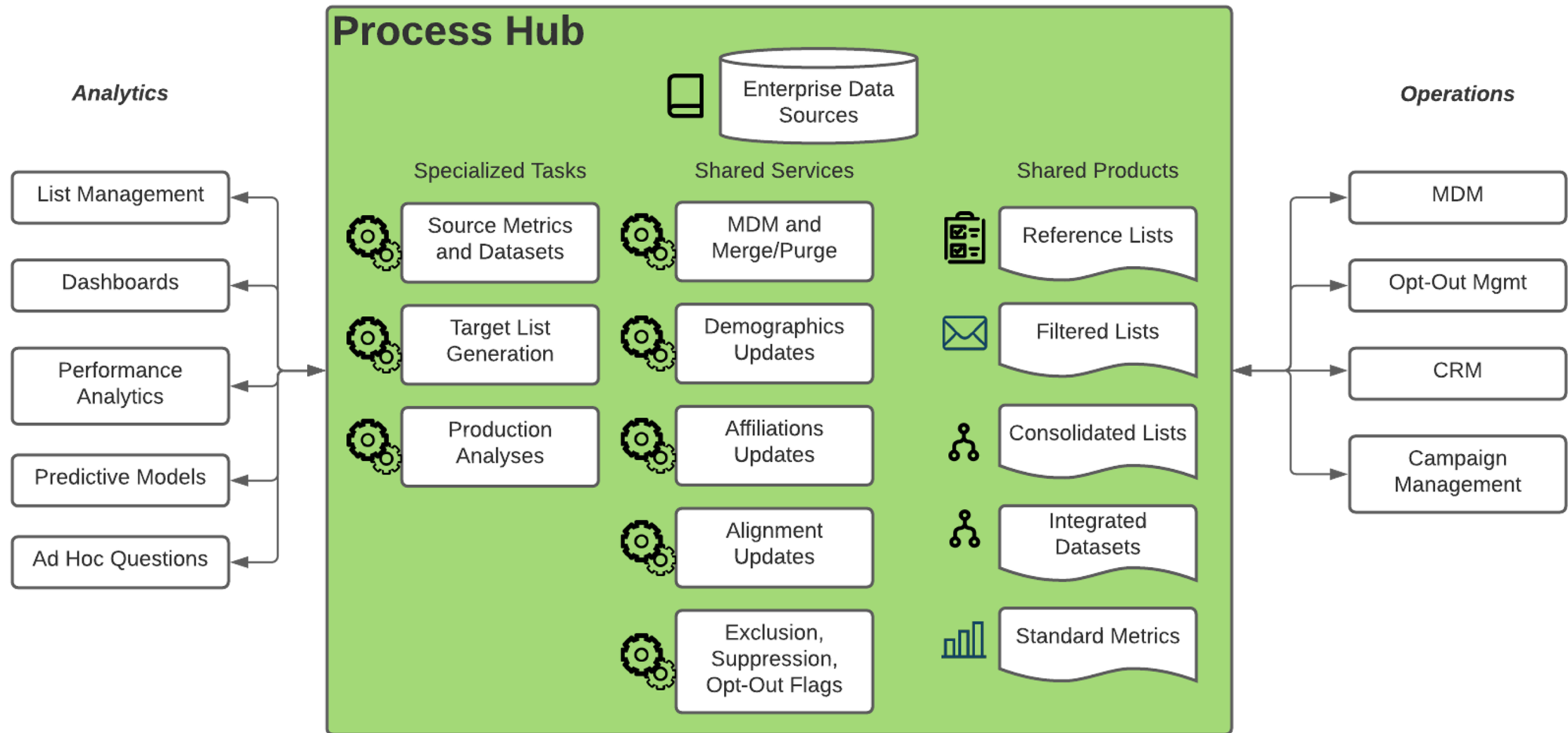
Process Hub in action lowering the cost-per-question by:

- Integrated code, data in a collaborative process hub.
- Multi-company builders, users & operators.

Benefit:

- Automated, governed & reusable processes to minimize consulting.
- Multi-team, multi-vendor, multi-location collaboration saves time.
- Open, secure platform for anyone to
 - Access data
 - Change the process used to create that data.

If you centralize the process, you can generalize the process to save time & effort

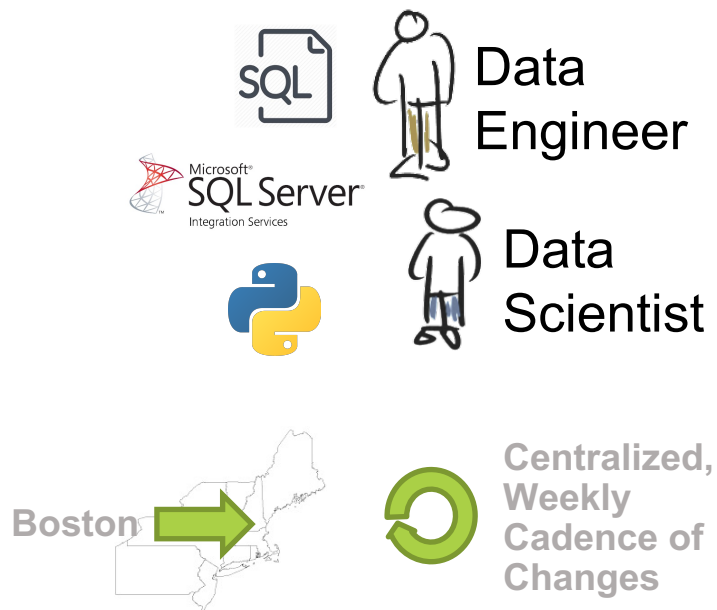


DataKitchen DataOps Process Hub Agenda

- Current Situation of Business Analytics (BA) Teams
- DataKitchen DataOps Process Hub for BA Teams
- Examples of BA Teams Using DataKitchen's DataOps Process Hub
- DataKitchen DataOps Process Hub Demo Example
- Conclusion & Questions

Coordinating Two Teams, Two Locations & Multiple Tools

Home Office Team



Local Office 'Self-Service' Team



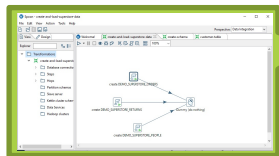
Challenges with Coordination

Home Office Team

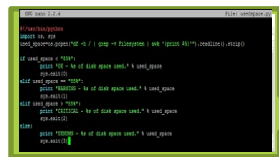
Make a change in schema?

Not Available for All?

New Data & Schema



Data Engineer



Data Scientist

Local Office Team

Break Reports?

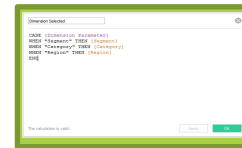
Add New Data Sets

Change Report Calculations

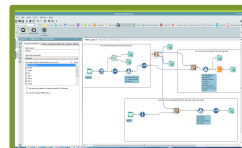
Inconsistencies?

Update/New Report

Not Working?



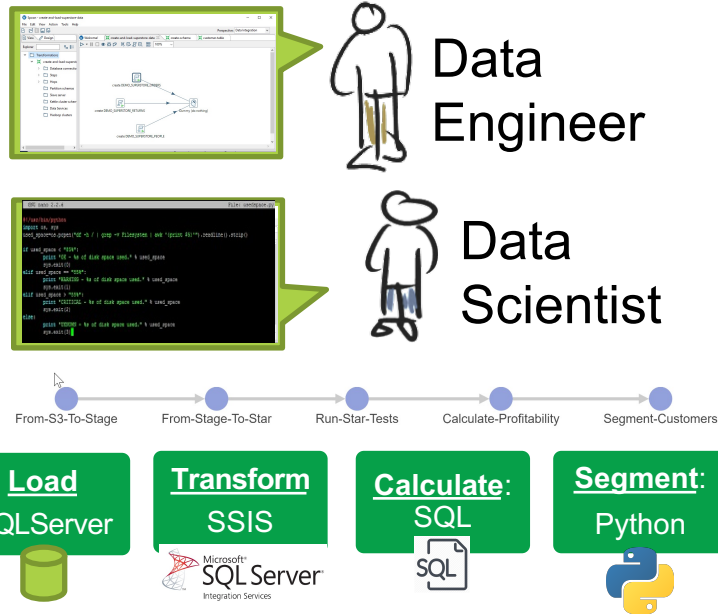
Data Analyst



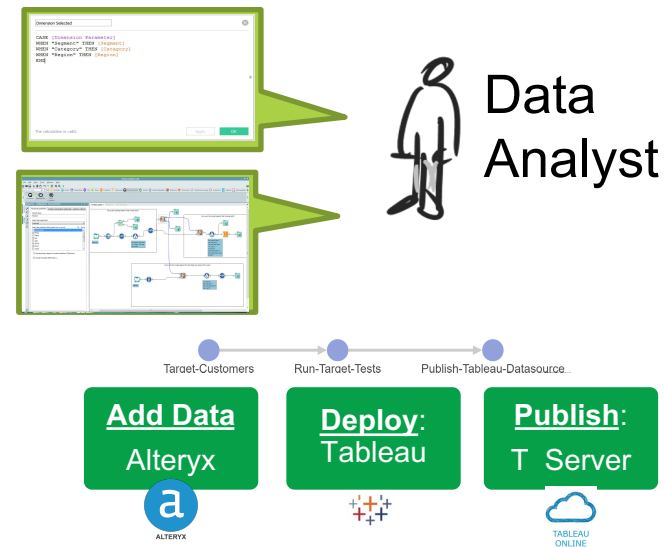
VP Marketing

Shared Result, Separate Responsibilities

Home Office Team

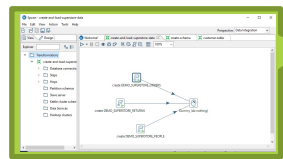


Local Office 'Self-Service' Team

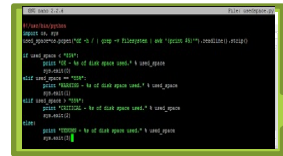


Overall Orchestration

Home Office Team

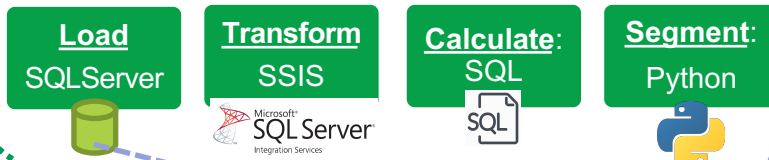


Data Engineer

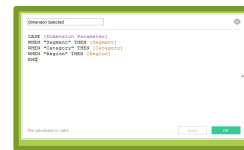


Data Scientist

From-S3-To-Stage → From-Stage-To-Star → Run-Star-Tests → Calculate-Profitability → Segment-Customers



Local Office 'Self-Service' Team



Data Analyst

Target-Customers → Run-Target-Tests → Publish-Tableau-Datasource...



VP Marketing



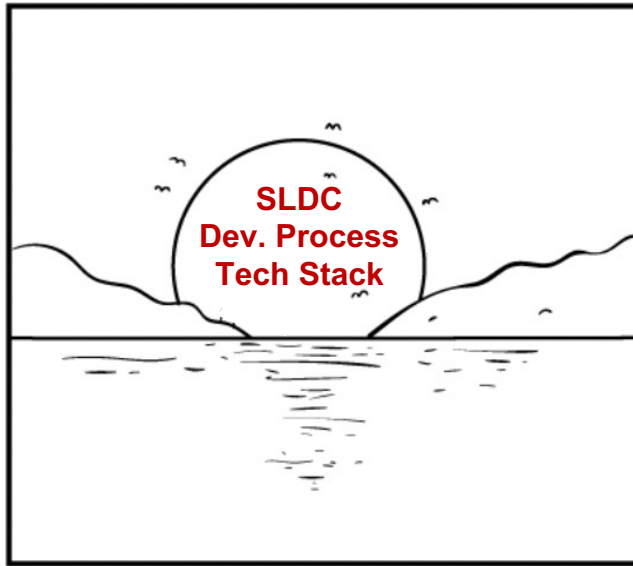
run-home-ingredient

run-local-ingredient

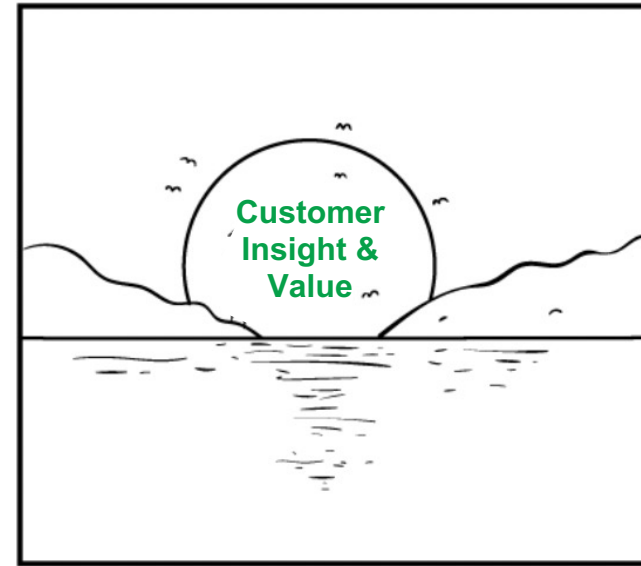
DataKitchen DataOps Process Hub Agenda

- Current Situation of Business Analytics (BA) Teams
- DataKitchen DataOps Process Hub for BA Teams
- Examples of BA Teams Using DataKitchen's DataOps Process Hub
- DataKitchen DataOps Process Hub Demo Example
- Conclusion & Questions

Two Different 'Suns' in Data & Analytic Teams

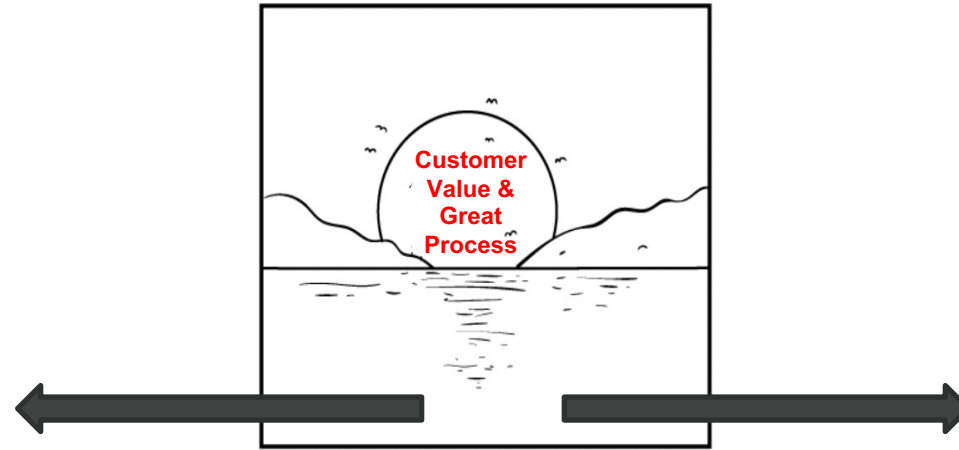


- For **many IT Teams**, the 'sun' that they first & foremost work toward is their development process, their 'SLDC', &/or following their tech standards
- End Customer Value & Insight come second



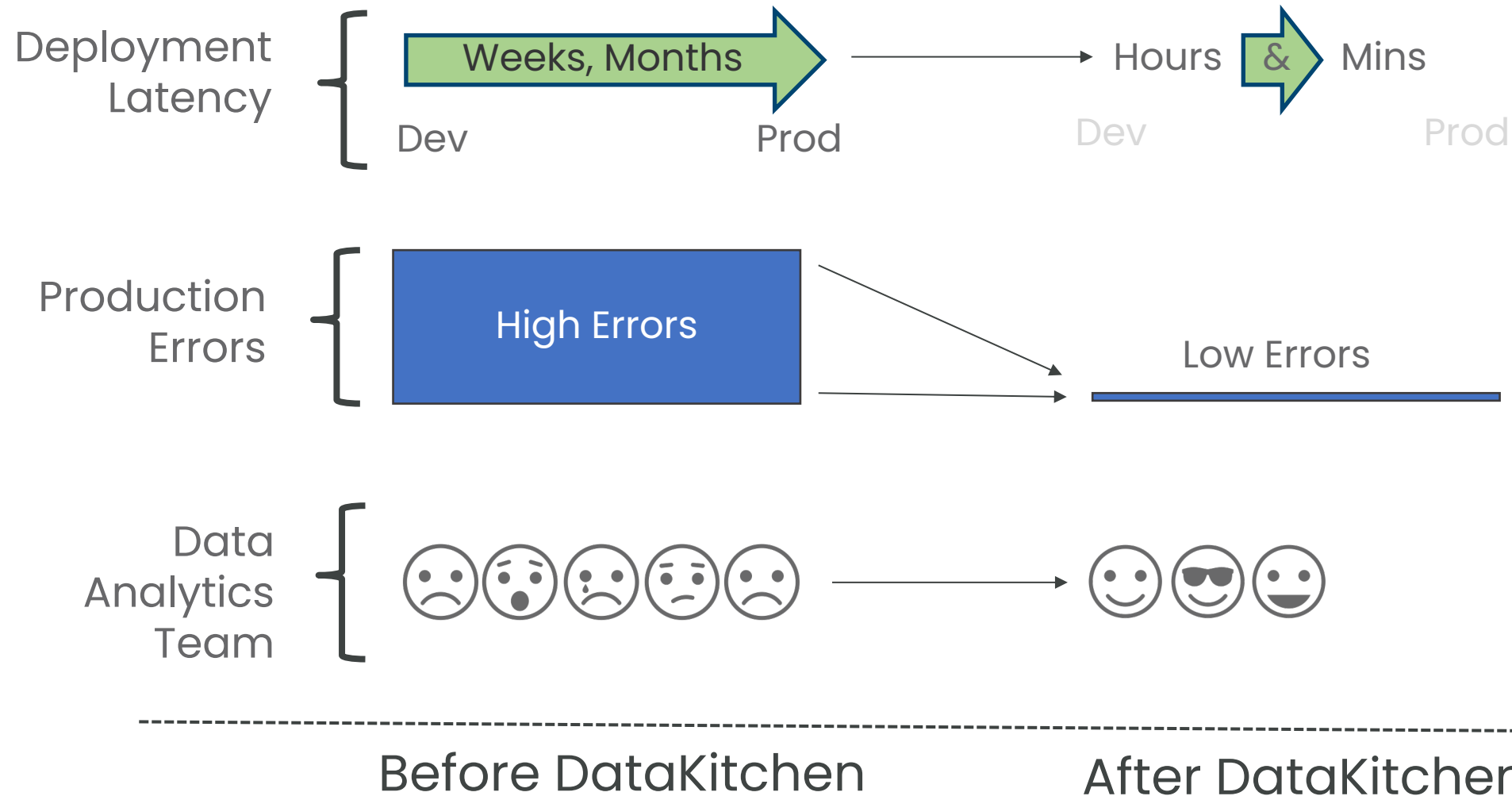
- For many **Business Analytic Teams**, the 'sun' that they first & foremost work toward is customer insight & making valuable analytics
- Repeated processes & scale come second

The Process Hub Allows Everyone to Win



- Robust process
 - Provides a win for IT
 - Improves IT Governance
 - Anyone (including IT) can leverage
 - Automation drives cost reductions
 - A single view of business logic
 - Transparent, open platform
 - Utilize existing data infrastructure
- Fast customer value
 - Lowers the cost-per-question
 - Controlled & trusted by Business Analytics
 - Reduces time to do repetitive production
 - Speeds the delivery of ad hoc analysis
 - Works with IT data
 - Minimizes errors & repeated steps
 - Allows multiple partners to leverage

Process Hub Benefit: Faster, Better & Happier



Learn More About DataOps & DataKitchen

Next Steps:

1. Sign [The DataOps Manifesto](#)
2. Read [The DataOps Cookbook](#)
3. Read [Recipes for DataOps Success](#)
4. Take [The DataOps Maturity Model Assessment](#)



Conclusion & Questions

- The DataKitchen DataOps Process Hub solves the 'last-mile' problem for Business Analytics Teams
- Questions?