

The this deck and its contents are the intellectual property of

#### Ben Biddle

made available for use under a Creative Common License Attribution-NonCommercial-ShareAlike 4.0 International (<a href="CC BY-NC-SA 4.0">CC BY-NC-SA 4.0</a>)

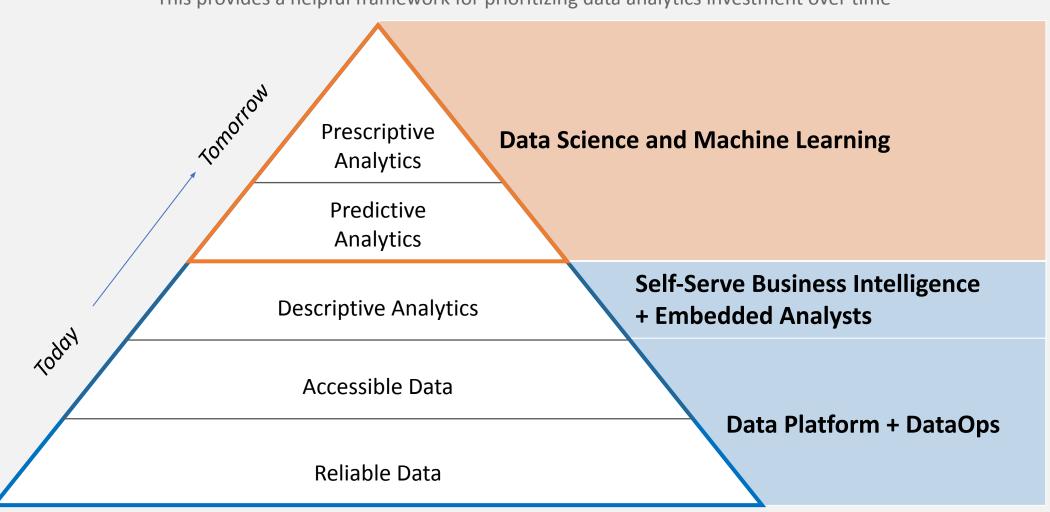
# A best-in-class data analytics org starts with a clear purpose

Help business stakeholders to make better decisions with reliable information and insights

- Data is just information in a digital format
- Insights come from discerning the patterns in the data
- Reliability implies both <u>accuracy</u> and <u>accessibility</u>
- Reliability does not imply certainty; the goal is to reduce uncertainty
- The value of your information and insights can be measured by the realized improvements to your decision making, e.g. better decisions made faster

# In pursuit of that purpose, data-and-analytics needs tend to follow a hierarchy similar to Maslow's

This provides a helpful framework for prioritizing data analytics investment over time



more cost effectively

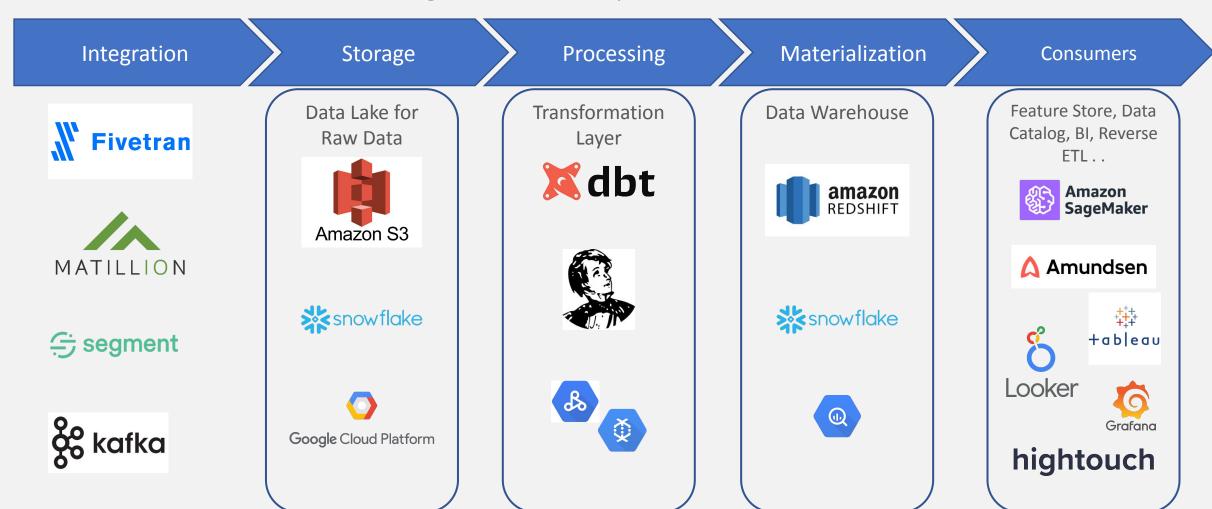
# By investing in the right tooling (i.e. Data Platform), your data-and-analytics team can scale more efficiently

Materialization **Processing** Integration Storage Consumers Use 3<sup>rd</sup> party Enable rapid Adopt a data Establish centrally Establish a single tools, such as pane of glass for exploratory modeling tool, managed trusted sources for slow Fivetran, to analysis using such as dbt, to business intelligence changing, frequently and launch a data integrate siloed replicated raw manage data sets sources in the data-as-code in used data sets in the catalog for discovery data lake Git data warehouse Consider replacing Consider adopting a star schema or Add more Empower users at 3<sup>rd</sup> party solutions enterprise knowledge graph to make the the periphery to consumers, such as data lake/warehouse easier to navigate compile their own with open source or a reverse ETL internally developed novel data sets solution, to the tools that can scale under a less-trusted stack

governance tier

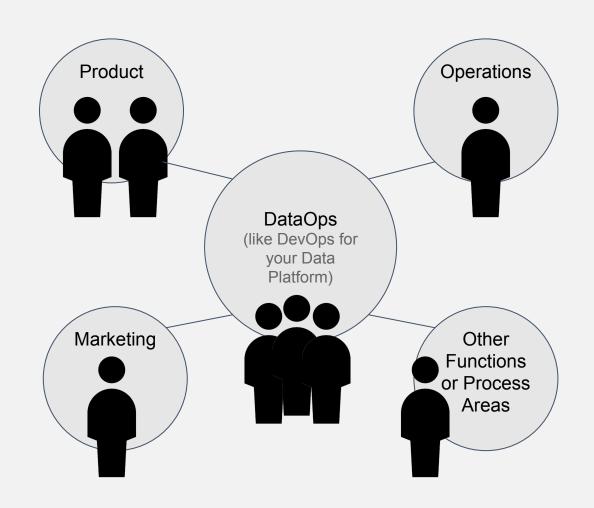
#### The Data Platform should enable you to ingest, transform and consume data from diverse sources and for distinct purposes

Data integrations will be important in the short term to get everyone at the company visibility into the same information about the growth, health and performance of the business



# Don't buy into the false dichotomy of centralized vs embedded data analytics

- Treat your Data Platform like an actual product, with a centralized "DataOps" team supporting the Data Platform
- Embed analysts with stakeholders where they can form stronger working relationships, deepen their business understanding, and acquire greater data fluency
- Roll both centralized and embedded subteams up under the same single point of accountability to improve coordination
- Make that single point of accountability independent of the stakeholders they serve (e.g. under the CTO, CFO or even CDO, not Product or Marketing)



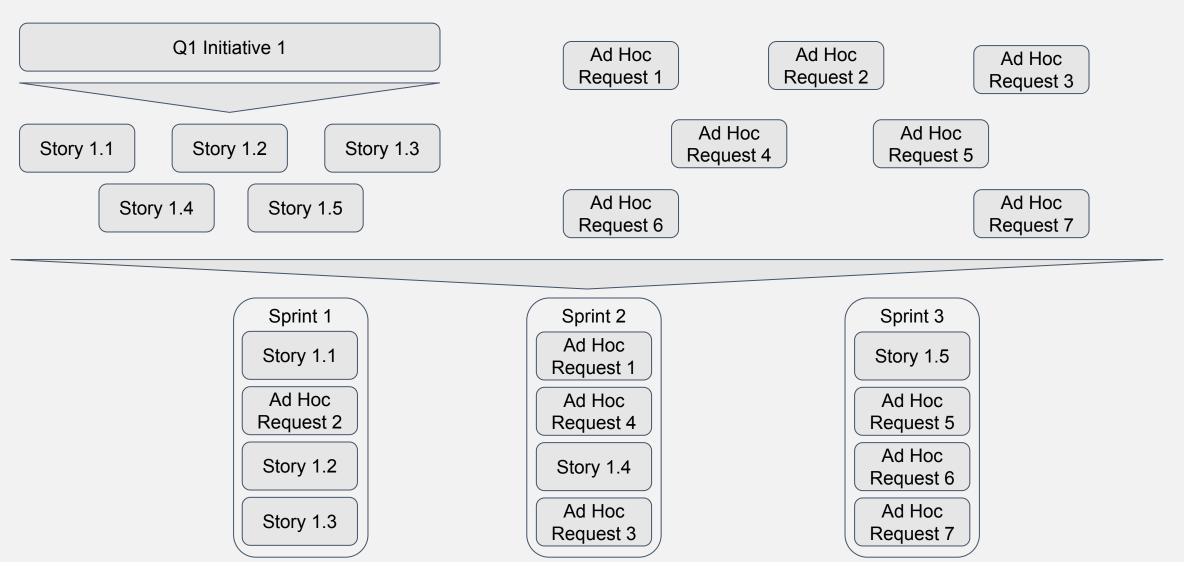
#### Grow headcount slowly and prioritize tooling and infrastructure investments to set new hires up for success

- Guiding principle of tech investments should be reducing complexity to increase efficiency, productivity and scalability; operationally, reduce uncertainty to improve decision velocity (both speed and direction)
- If using outside partners to set up tooling, ensure there is a seamless hand-off to the teams that will have to support those tools on an ongoing basis
- Hire analysts who are experts in SQL and business intelligence tools with the stakeholder management skills necessary to navigate competing business priorities
- Start with generalists and add in more specialization over time to serve higher level needs in the hierarchy; align analysts to business units and/or process areas to maximize productivity (e.g. closer relationships, better data fluency, deeper domain understanding)
- Use a tiered data governance model to balance speed and agility with quality and reliability, with the governance tier reflecting the use cases and level of trust in the data

Use quarterly roadmaps together with agile sprints to keep the focus on what's important and not just what's urgent

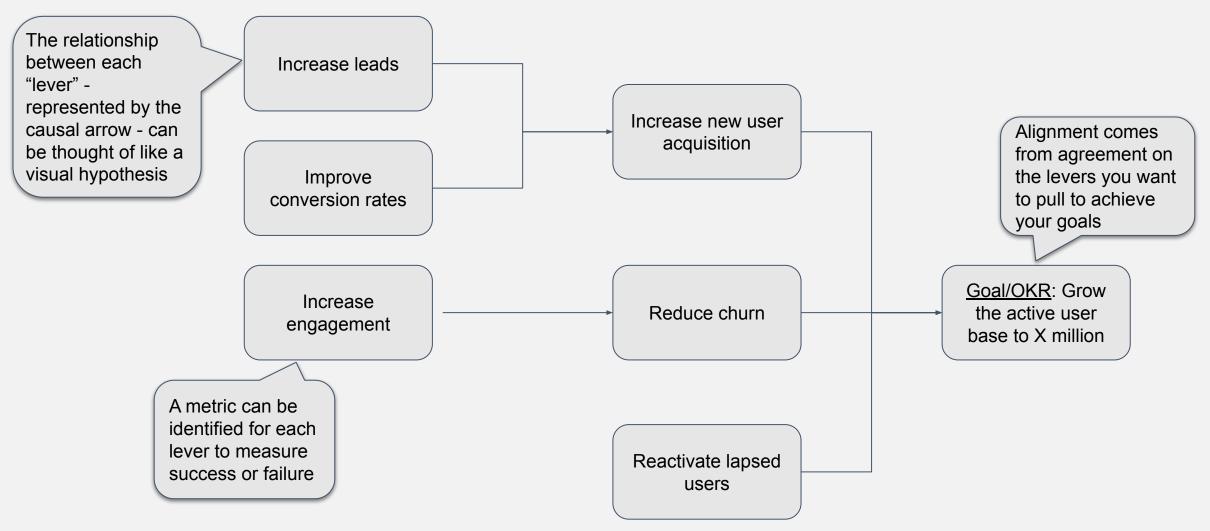


# As each quarter evolves, ad hoc requests can be incorporated into sprint planning to stay responsive to changing needs



## To ensure key performance indicators and operating metrics are also aligned with company goals and OKRs, causality is key

The diagram of causal relationships below is purely illustrative and should be considered complete or comprehensive



## Another helpful framework for identifying important metrics is the hypothesis dashboard described in <u>Getting to Plan B</u>

A DASHBOARD FOR JOHNNY'S LEMONADE STAND					
Hypotheses	Metrics	Actual period 1	Actual period 2	Actual period 3	Insights obtained, course corrections needed
Leap of faith 1: Commuters will stop and buy a refreshing drink					
Hypothesis 1: At least 10 customers per day	Customer count	2 customers	No one stopped in the rain	6 customers	High pricing deters sales, they look, don't buy; no point in setting up if it rains; seems like demand is somewhat less than Johnny thought.
Leap of faith 2: People will pay a premium price					
Hypothesis 2: \$1.50 per glass will be accept- able	Total sales, price paid	\$3.00 total sales, \$1.50 per glass		\$5.50 in sales (1@ 50 cents, 5@ \$1)	\$1.50 too high, based on Monday sales; pricing then reduced; \$1 looks about right based on Wednesday's lower pricing.

Image from <a href="http://getting-to-plan-b.com/Documents/SloanPlanB">http://getting-to-plan-b.com/Documents/SloanPlanB</a> PDF April2010.pdf