Introduction

Doug Gray

Methodologies

- Four pillars
- Five stages
- DELTTA
- FACE

DataScience@SMU

Introduction to Target Case Study

Doug Gray

Target Stores:
The Case of the Pregnant Teenager

- http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/
- https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&_r=1&hp

DataScience@SMU

Discussion of Target Case Study | DELTTAA

Doug Gray

DELTTAA **D**ata

Data:

- Guest ID: credit card bucket of history of what they have bought at Target

  The desire to collect information on customers is not new for target or any other large retailer, of course. **For decades, Target has collected vast amounts of data on every person who regularly walks into one of its stores. Whenever possible, Target assigns each shopper a unique code—known internally as the Guest ID number—that keeps tabs on everything they buy. "If you use a credit card or a coupon, or fill out a survey, or mail in a refund, or call the customer help line, or open an e-mail we've sent you or visit our Web site, we'll record it and link it to your Guest ID," Pole said. "We want to know everything we can."**

DELTTAA **D**ata

Data:

- Buying purchasing and demographic information from other sources

  Also linked to your Guest ID is demographic information like your age, whether you are married and have kids, which part of town you live in, how long it takes you to drive to the store, your estimated salary, whether you've moved recently, what credit cards you carry in your wallet and what Web sites you visit. Target can buy data about your ethnicity, job history, the magazines you read, if you've ever declared bankruptcy or got divorced, the year you bought (or lost) your house, where you went to college, what kinds of topics you talk about online, whether you prefer certain brands of coffee, paper towels, cereal, or applesauce, your political leanings, reading habits, charitable giving, and the number of cars you own. (In a statement, Target declined to identify what demographic information it collects or purchases.)

DELTTAA **D**ata

Data:

- Baby registry data

  The only problem is that identifying pregnant customers is harder than it sounds. **Target has a baby-shower registry, and Pole started there,** observing how shopping habits changed as a woman approached her due date, which women on the registry had willingly disclosed. He ran test after test, analysing the data, and before long some useful patterns emerged.

D**E**LTTAA **E**nterprise

Almost every major retailer, from grocery chains to investment banks to the U.S. Postal Service, has a "predictive analytics" department devoted to understanding not just consumers' shopping habits but also their personal habits, so as to more efficiently market to them. "But Target has always been one of the smartest at this," says Eric Siegel, a consultant and the chairman of a conference called Predictive Analytics World. "We're living through a golden age of behavioral research. It's amazing how much we can figure out about how people think now."

DE**L**TTAA **L**eadership

Soon after the new ad campaign began, Target's Mom and Baby sales exploded. The company doesn't break out figures for specific divisions, but between 2002—when Pole was hired—and 2010, Target's revenues grew from $44 billion to $67 billion. **In 2005, the company's president, Gregg Steinhafel, boasted to a room of investors about the company's "heightened focus on items and categories that appeal to specific guest segments such as mom and baby."**

DE**L**TTAA **L**eadership

Pole applied his program to every regular female shopper in Target's national database and soon had a list of tens of thousands of women who were most likely pregnant. If they could entice those women or their husbands to visit Target and buy baby-related products, the company's cue-routine-reward calculators could kick in and start pushing them to buy groceries, bathing suits, toys and clothing, as well. When Pole shared his list with the marketers, he said, they were ecstatic. **Soon, Pole was getting invited to meetings above his paygrade. Eventually his paygrade went up.**

Eventually, he became more of the leadership!

DELTTAA Targets

There are, however, some brief periods in a person's life when old routines fall apart and buying habits are suddenly in flux. One of those moments—*the* moment, really—is right around the birth of a child, when parents are exhausted and overwhelmed and their shopping patterns and brand loyalties are up for grabs. But as Target's marketers explained to Pole, timing is everything. Because birth records are usually public, the moment a couple have a new baby, they are almost instantaneously barraged with offers and incentives and advertisements from all sorts of companies. Which means that the key is to reach them earlier, before any other retailers know a baby is on the way. **Specifically, the marketers said they wanted to send specially designed ads to women in their second trimester, which is when most expectant mothers begin buying all sorts of new things, like prenatal vitamins and maternity clothing. "Can you give us a list?" the marketers asked.**

In other words:

1. Can you identify women that are pregnant before they deliver?
2. Can you identify women in the second trimester?

DELTTAA Technologies

Lead Data Scientist - Marketing/Ad Tech

3.4 ★ **Target** – Sunnyvale, CA

glassdoor

Proficiency in data
analytics languages such as SQL, R, Python.

Proficiency in
big data technologies such as Hadoop, Hive, Spark

Card scanning infrastructure for “Guest ID” tracking.

Search “target data scientist” in glassdoor: https://www.glassdoor.com/index.htm

DELTT**A**A **A**nalytical Techniques

Lead Data Scientist - Marketing/Ad Tech

3.4 ★ **Target** – Sunnyvale, CA

glassdoor

3+ years of experience
in large scale A/B and multivariate testing, or other relevant work in design of experiments.

Expertise in either
statistical modeling or machine learning (deep learning).

Conceivable techniques:
Clustering for customer segmentation

Search “target data scientist” in glassdoor: https://www.glassdoor.com/index.htm

DELTTA**A A**nalysts

Andrew Pole had just started working as a statistician for Target in 2002, when two colleagues from the marketing department stopped by his desk to ask an odd question: “If we wanted to figure out if a customer is pregnant, even if she didn’t want us to know, can you do that?”

DELTTA**A A**nalysts

Pole has a master's degree in statistics and another in economics, and has been obsessed with the intersection of data and human behaviour most of his life. His parents were teachers in North Dakota, and while other kids were going to 4-H, Pole was doing algebra and writing computer programs. "The stereotype of a math nerd is true," he told me when I spoke with him last year. "I kind of like going out and evangelizing analytics."

DELTTA**A A**nalysts

Andrew Pole • 3rd

Vice President - Workforce, Multicultural, and Personalization Analytics at U.S. Bank

Greater Minneapolis-St. Paul Area

Connect ...

U.S. Bank

Kansas State University

See contact info

500+ connections

DELTTA**A Analysts**

Andrew Pole • 3rd

Vice President - Workforce, Multicultural, and Personalization Analytics at U.S. Bank

Greater Minneapolis-St. Paul Area

Connect ...

U.S. Bank

Kansas State University

See contact info

500+ connections

Experience

Vice President - Workforce, Multicultural, and Personalization Analytics
U.S. Bank
Jun 2018 – Present · 8 mos
Minneapolis, MN

Lead workforce, multicultural, and personalization analytics at US Bank

Target
17 yrs

Sr. Group Manager - Marketing BI
Nov 2011 – Present · 7 yrs 3 mos
Minneapolis, MN

Manage a team of approximately 30 analysts in the US and 30 analysts in India in support of direct and online marketing analytics including campaign design and execution, guest segmentation, predictive modeling, ad hoc analysis, campaign reporting, and data governance

Director, Guest Data Management
Nov 2002 – Present · 16 yrs 3 mos

Manage and improve foundational guest (customer) data for the Target enterprise

Senior Manager, Guest Analytics
2002 – Nov 2011 · 9 yrs

Managed a team of approximately 30 analysts in the US and 30 analysts in India in support of direct and online marketing analytics including campaign design and execution, guest segmentation, predictive modeling, ad hoc analysis, and campaign reporting
Led direct and online analytical teams responsible for marketing programs that gen... See more

Show fewer roles

DataScience@SMU

Discussion of Target Case Study | FACE

Doug Gray

FACE

- Framing the problem
  - Problem recognition
  - Review of previous findings
  - Modeling approach
- Solving the problem
  - Data collection
  - Modeling
  - Data analysis
- Communicating and acting on results
  - Results presentation and action, i.e., “telling a story”
  - Creating impactful visualization
- Embedding final models and methods in enterprise business processes and systems

FACE

Framing the problem

- Problem recognition

  Andrew Pole had just started working as a statistician for Target in 2002, when two colleagues from the marketing department stopped by his desk to ask an odd question: **“If we wanted to figure out if a customer is pregnant, even if she didn’t want us to know, can you do that?”**

FACE

Framing the problem

- Problem recognition:

  Andrew Pole was hired by Target to use the same kinds of insights into consumers' habits to expand Target's sales. His assignment was to analyse all the cue-routine-reward loops among shoppers and help the company figure out how to exploit them. Much of his department's work was straightforward: find the customers who have children and send them catalogs that feature toys before Christmas. Look for shoppers who habitually purchase swimsuits in April and send them coupons for sunscreen in July and diet books in December. **But Pole's most important assignment was to identify those unique moments in consumers' lives when their shopping habits become particularly flexible and the right advertisement or coupon would cause them to begin spending in new ways. And among life events, none are more important than the arrival of a baby. At that moment, new parents' habits are more flexible than at almost any other time in their adult lives. If companies can identify pregnant shoppers, they can earn millions.**

FACE

Framing the problem

- Problem recognition:
- Review of previous findings

> **The reason Target can snoop on our shopping habits is that, over the past two decades, the science of habit formation has become a major field of research in neurology and psychology departments at hundreds of major medical centers and universities, as well as inside extremely well financed corporate labs.** "It's like an arms race to hire statisticians nowadays," said Andreas Weigend, the former chief scientist at Amazon.com.
>
> **…**
>
> exist. **One study from Duke University estimated that habits, rather than conscious decision-making, shape 45 percent of the choices we make every day, and recent discoveries have begun to change everything from the way we think about dieting to how doctors conceive treatments for anxiety, depression, and addictions.**

FACE

Framing the problem

- Problem recognition:
- Review of previous findings
- Modeling approach

F**A**CE

Solving the problem

- Data collection

  The only problem is that identifying pregnant customers is harder than it sounds. **Target has a baby-shower registry, and Pole started there, observing how shopping habits changed as a woman approached her due date, which women on the registry had willingly disclosed.** He ran test after test, analysing the data, and before long some useful patterns emerged.

F**A**CE

Solving the problem

- Data collection

The desire to collect information on customers is not new for Target or any other large retailer, of course. **For decades, Target has collected vast amounts of data on every person who regularly walks into one of its stores. Whenever possible, Target assigns each shopper a unique code—known internally as the Guest ID number—that keeps tabs on everything they buy.** "If you use a credit card or a coupon, or fill out a survey, or mail in a refund, or call the customer help line, or open an e-mail we've sent you or visit our Web site, we'll record it and link it to your Guest ID," Pole said. "We want to know everything we can."

F**A**CE

Solving the problem

- Data collection

  Also linked to your Guest ID is demographic information like your age, whether you are married and have kids, which part of town you live in, how long it takes you to drive to the store, your estimated salary, whether you've moved recently, what credit cards you carry in your wallet and what web sites you visit. Target can buy data about your ethnicity, job history, the magazines you read, if you've ever declared bankruptcy or got divorced, the year you bought (or lost) your house, where you went to college, what kinds of topics you talk about online, whether you prefer certain brands of coffee, paper towels, cereal or applesauce, your political leanings, reading habits, charitable giving and the number of cars you own. (In a statement, Target declined to identify what demographic information it collects or purchases.)

F**A**CE

Solving the problem

- Data collection
- Modeling

Andrew Pole had just started working as a statistician for Target in 2002, when two colleagues from the marketing department stopped by his desk to ask an odd question: “If we wanted to figure out if a customer is pregnant, even if she didn’t want us to know, can you do that?”

Predicting Pregnant / Not Pregnant: Logistic Regression, SVM, LDA, Latent Class Modeling, Random Forest, CHAID, KNN

F**A**CE

Solving the problem

- Data collection
- Modeling

As Pole's computers crawled through the data, he was able to identify about 25 products that, when analysed together, allowed him to assign each shopper a "pregnancy prediction" score. More important, he could also estimate her due date to within a small window, so Target could send coupons timed to very specific stages of her pregnancy.

F**A**CE

Solving the problem

- Data collection
- Modeling
- Data analysis

date, which women on the registry had willingly disclosed. **He ran test after test, analysing the data, and before long some useful patterns emerged.** Lotions, for example. Lots of people buy lotion, but one of Pole's colleagues noticed that women on the baby registry were buying larger quantities of unscented lotion around the beginning of their second trimester. **Another analyst noted that sometime in the first 20 weeks, pregnant women loaded up on supplements like calcium, magnesium and zinc. Many shoppers purchase soap and cotton balls, but when someone suddenly starts buying lots of scent-free soap and extra-big bags of cotton balls, in addition to hand sanitizers and washcloths, it signals they could be getting close to their delivery date.**

FACE

Communicating and acting on results

- Results presentation and action, i.e., “telling a story”
- Creating impactful visualization

- Strong communication skills to influence cross functional teams.

Lead Data Scientist - Marketing/Ad Tech

3.4 ★ **Target** – Sunnyvale, CA

Job

Requirements:

- Masters or PhD in statistics, applied mathematics, computer science, econometrics or an equivalent domain.
- 3+ years of experience in large scale A/B and multivariate testing, or other relevant work in design of experiments.
- Expertise in either statistical modeling or machine learning (deep learning).
- Proficiency in data analytics languages such as SQL, R, Python.
- Proficiency in big data technologies such as Hadoop, Hive, Spark.
- Strong communication skills to influence cross functional teams.

FACE

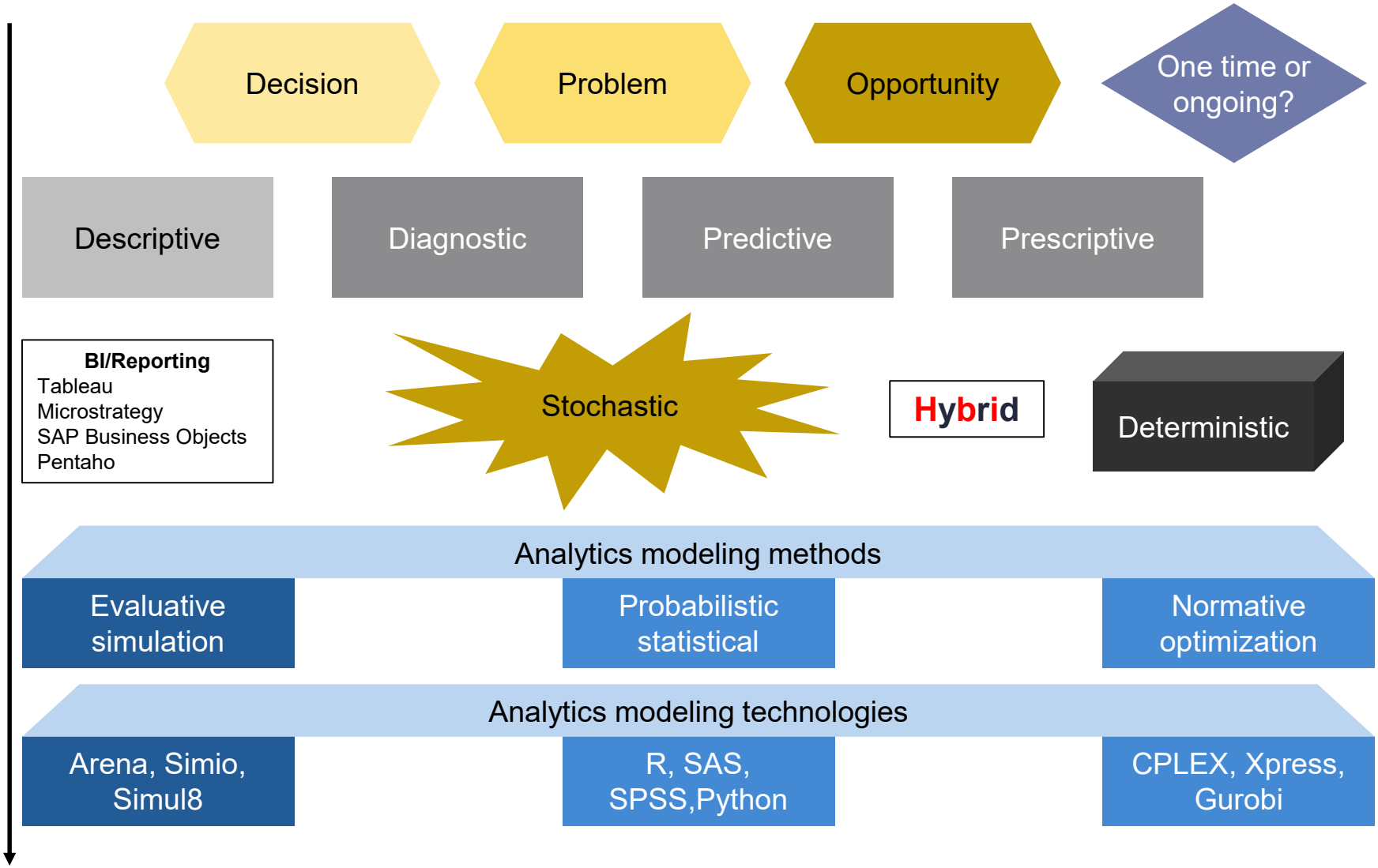
Embedding final models and methods in enterprise business processes and systems

know a baby is on the way. **Specifically, the marketers said they wanted to send specially designed ads to women in their second trimester, which is when most expectant mothers begin buying all sorts of new things, like prenatal vitamins and maternity clothing.** "Can you give us a list?" the marketers asked.

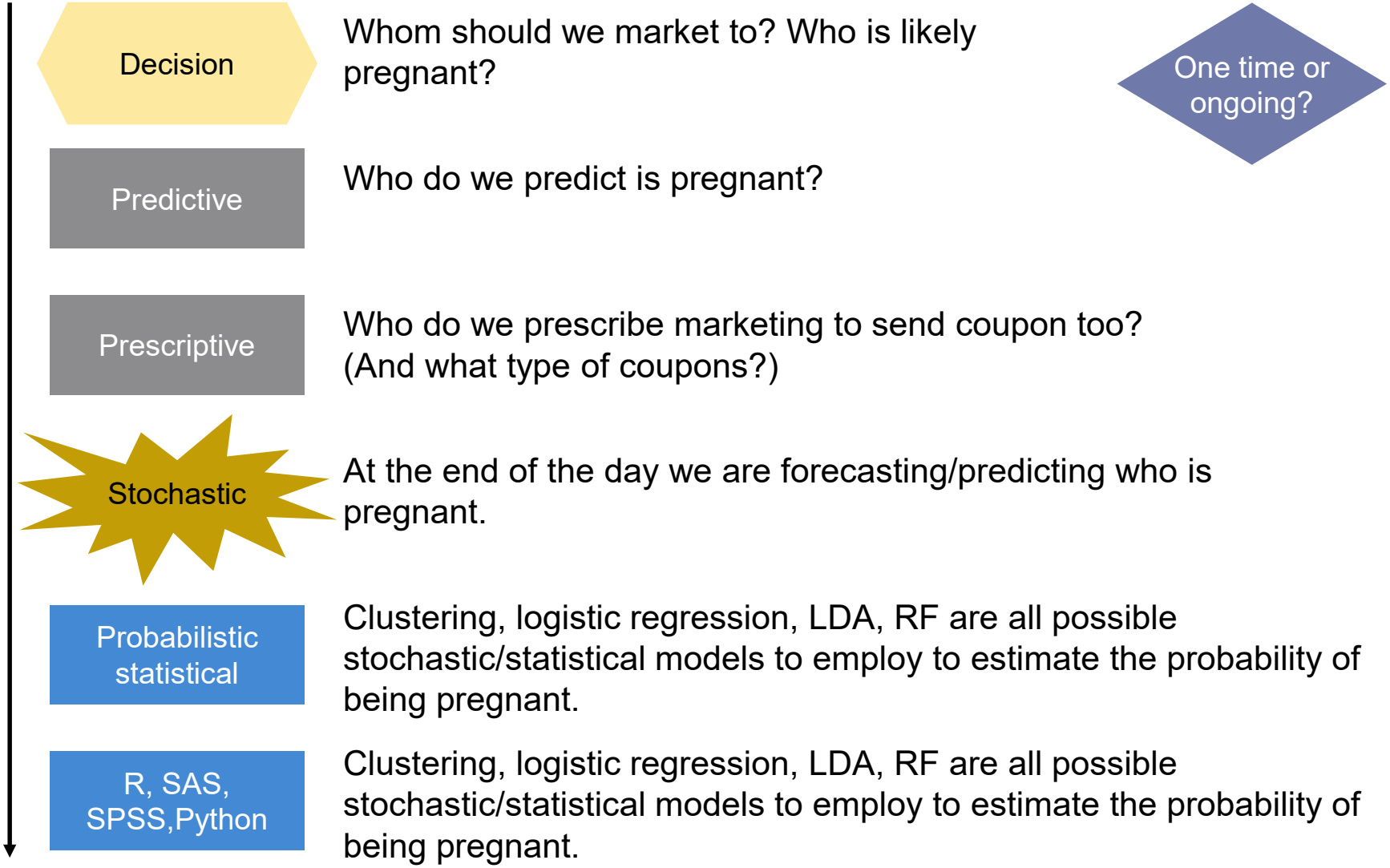
Pole applied his program to every regular female shopper in Target's national database and soon had a list of tens of thousands of women who were most likely pregnant. If they could entice those women or their husbands to visit Target and buy baby-related products, the company's cue- routine-reward calculators could kick in and start pushing them to buy groceries, bathing suits, toys and clothing, as well. **When Pole shared his list with the marketers, he said, they were ecstatic.** Soon, Pole was getting invited to meetings above his paygrade. Eventually his paygrade went up.

FACE: The Pachinko Machine!

Frame and Solve: Analytics Solution Approach



Frame and Solve: Analytics Solution Approach



DataScience@SMU

Introduction to Additional Case Studies

Doug Gray

DataScience@SMU

Additional Case Studies

Doug Gray

Uber / Lyft / Didi Chuxing

How do they know what to charge?

- https://www.fastcompany.com/company/uber
- https://www.investopedia.com/articles/personal-finance/111015/story-uber.asp
- https://www.forbes.com/sites/bizcarson/2018/09/19/where-uber-is-winning-the-world-and-where-it-has-lost/#62e92f0b4d6e
- https://www.forbes.com/sites/jinshanhong/2017/08/03/how-chinas-ride-hailing-king-didi-is-taking-over-the-world-before-uber-can/#2eab47cd3fd8
- https://www.forbes.com/sites/bernardmarr/2018/11/26/ai-in-china-how-uber-rival-didi-chuxing-uses-machine-learning-to-revolutionize-transportation/#3b7a64326732

Walmart

- Home of the 13 petabyte data warehouse
- Descriptive analytics on steroids

UPS ORION

- Road integrated optimization and navigation
- 120 deliveries per day each by 55,000 drivers
- 120 uniquely different routes to deliver them

https://www.ups.com/us/en/services/knowledge-center/article.page?name=orion-the-algorithm-proving-that-left-isn-t-right&kid=aa3710c2

https://www.pressroom.ups.com/pressroom/ContentDetailsViewer.page?ConceptType=Factsheets&id=1426321616277-282

Netflix

How many copies of ***Top Gun*** to buy and where to store them?

http://www.cnet.com/news/netflix-awards-1-million-for-outdoing-cinematch/

DataScience@SMU

Introduction to Amazon Case Study

Doug Gray

Amazon

How did they know you like *Tom Clancy* novels, and how do they get them to you?

http://www.wired.com/2015/10/get-used-to-amazon-being-a-profitable-company/

DataScience@SMU

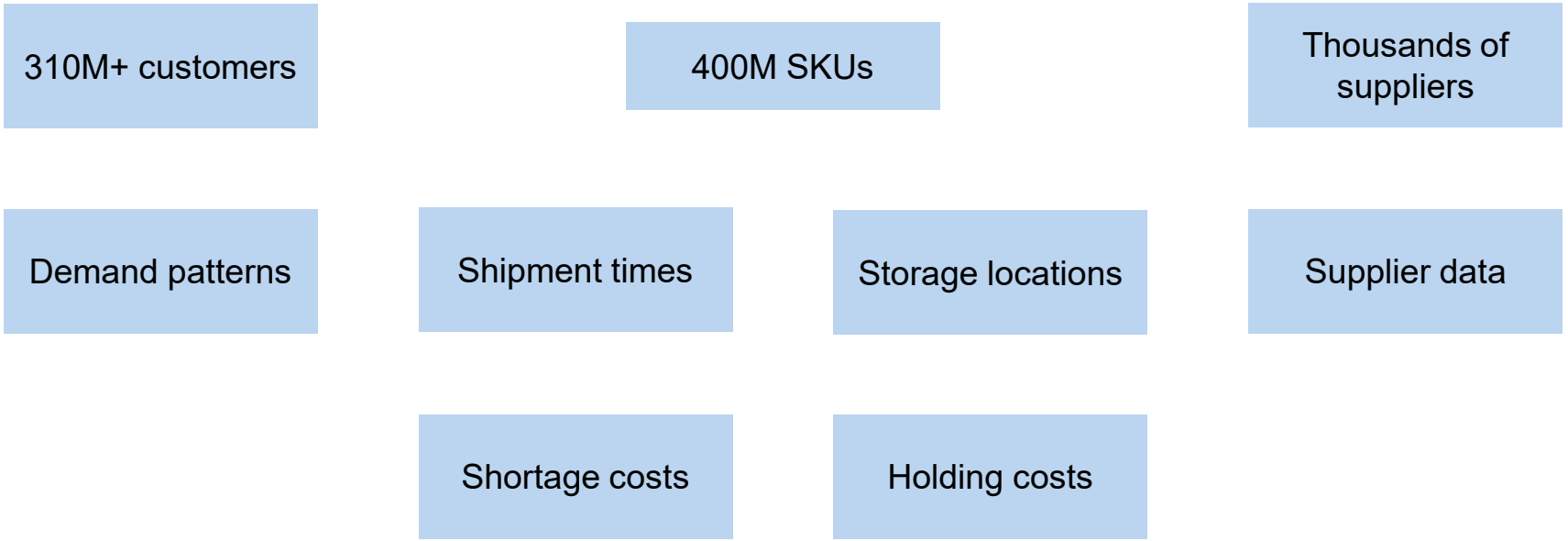
Discussion of Amazon Case Study | DELTTAA

Doug Gray

Amazon Retail **D**ELTTAA

"***We never throw any data away, regardless of whether or not we know what we are going to do with it.***"

— Jeff Bezos



Amazon Retail D**E**LTTAA

"Supply Chain Optimization Technologies (SCOT) is the core of Amazon Retail and has streamlined, automated, and optimized 100% of every aspect of our retail supply chain processes, decision-making enterprise-wide using predictive and prescriptive analytics"

310M+ customers

400M SKUs

Thousands of suppliers

How much to buy?

Shipment times

Storage locations

Where to store it?

Shortage costs

Holding costs

Amazon Retail DE**L**TTAA

"Supply chain optimization technologies (SCOT) is the core of Amazon Retail and has streamlined, automated, and optimized 100% of every aspect of our retail supply chain processes, decision-making enterprise-wide using predictive and prescriptive analytics."

Dr. Gang Yu
Wharton School, University of Pennsylvania
PhD, Operations Research
VP World Wide Supply Chain Operations

"We never throw any data away, regardless of whether or not we know what we are going to do with it."

— Jeff Bezos

Jeff Bezos
Princeton University
B.S. Electrical Engineering and Computer Science
Chairman and CEO

http://www.se.cuhk.edu.hk/yu-gang

Amazon Retail DEL**T**TAA

Inventory turns

Lead times

YOY sales growth

Gross margin

Targets

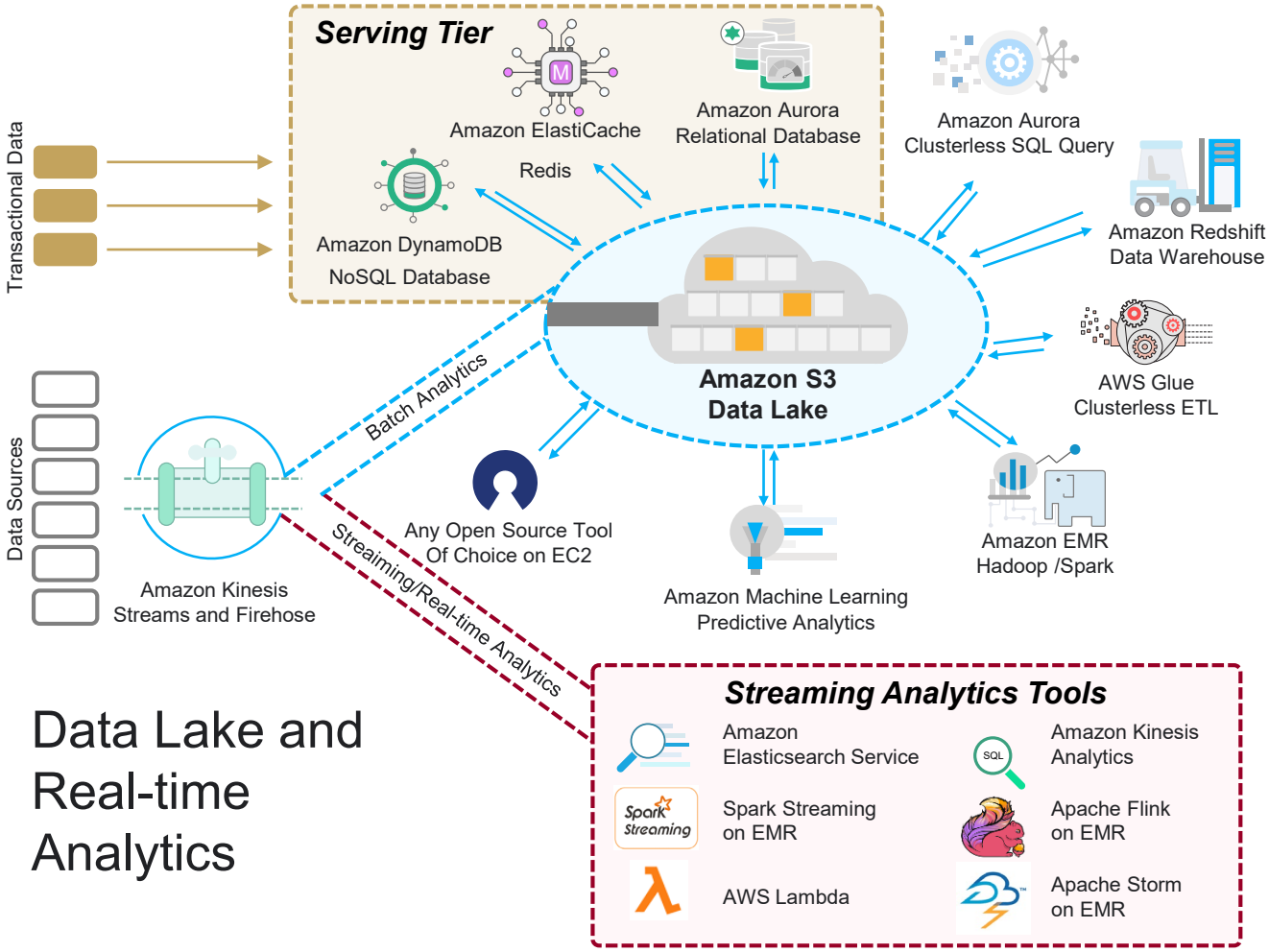
Delivery time performance

Shortage costs

Holding costs

Delivery costs

Amazon Retail
DELTTAA (Amazon Web Services AWS)



Data Lake and
Real-time
Analytics

Amazon Retail DELTT**A**A

Predictive		Prescriptive	
310M+ customers	400M SKUs		Thousands of suppliers
Demand patterns	Shipment times	Storage locations	Supplier data
How much to buy?	Shortage costs	Holding costs	Where to store it?

Amazon Retail DELTTA**A**

SCOT = 1,200 Omids

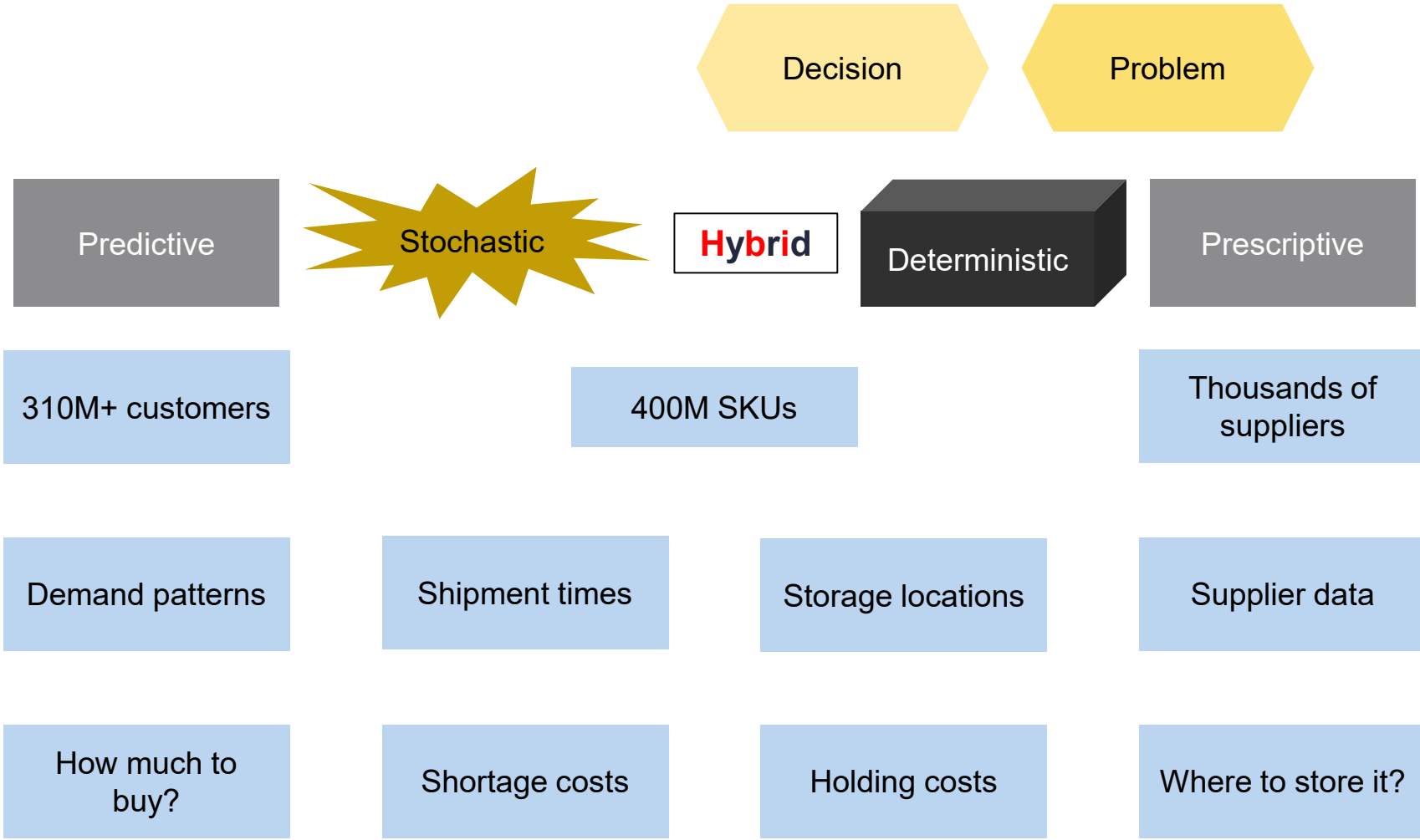
Dr. Alirezai “Omid” Madadi
Clemson University
PhD, Industrial Engineering
Thesis: Retail Supply Chain Optimization
Amazon SCOT Research Scientist II

DataScience@SMU

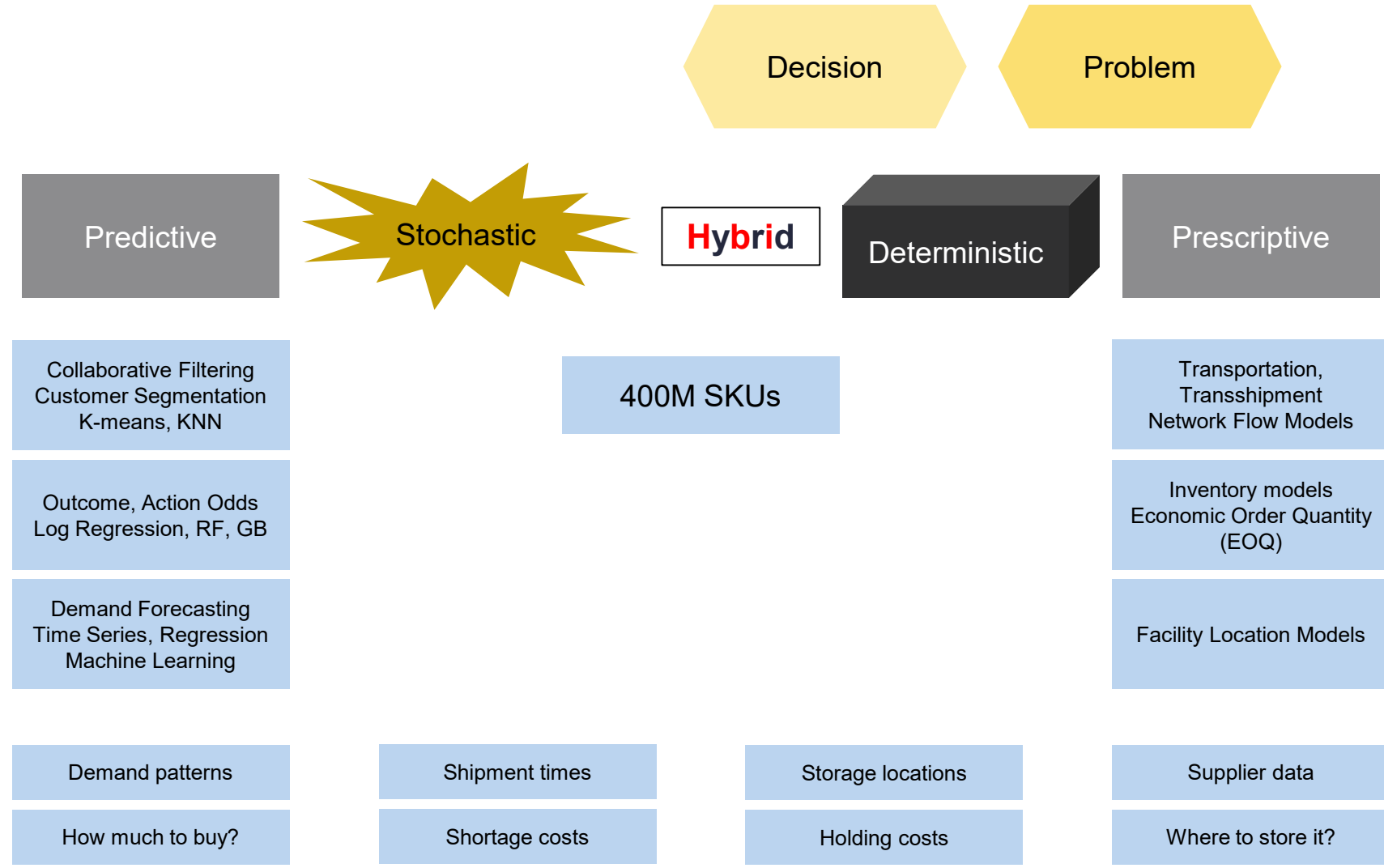
Discussion of Amazon Case Study | FACE

Doug Gray

Amazon Retail FACE

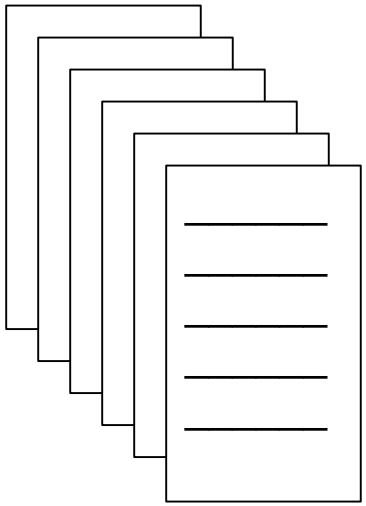


Amazon Retail FACE



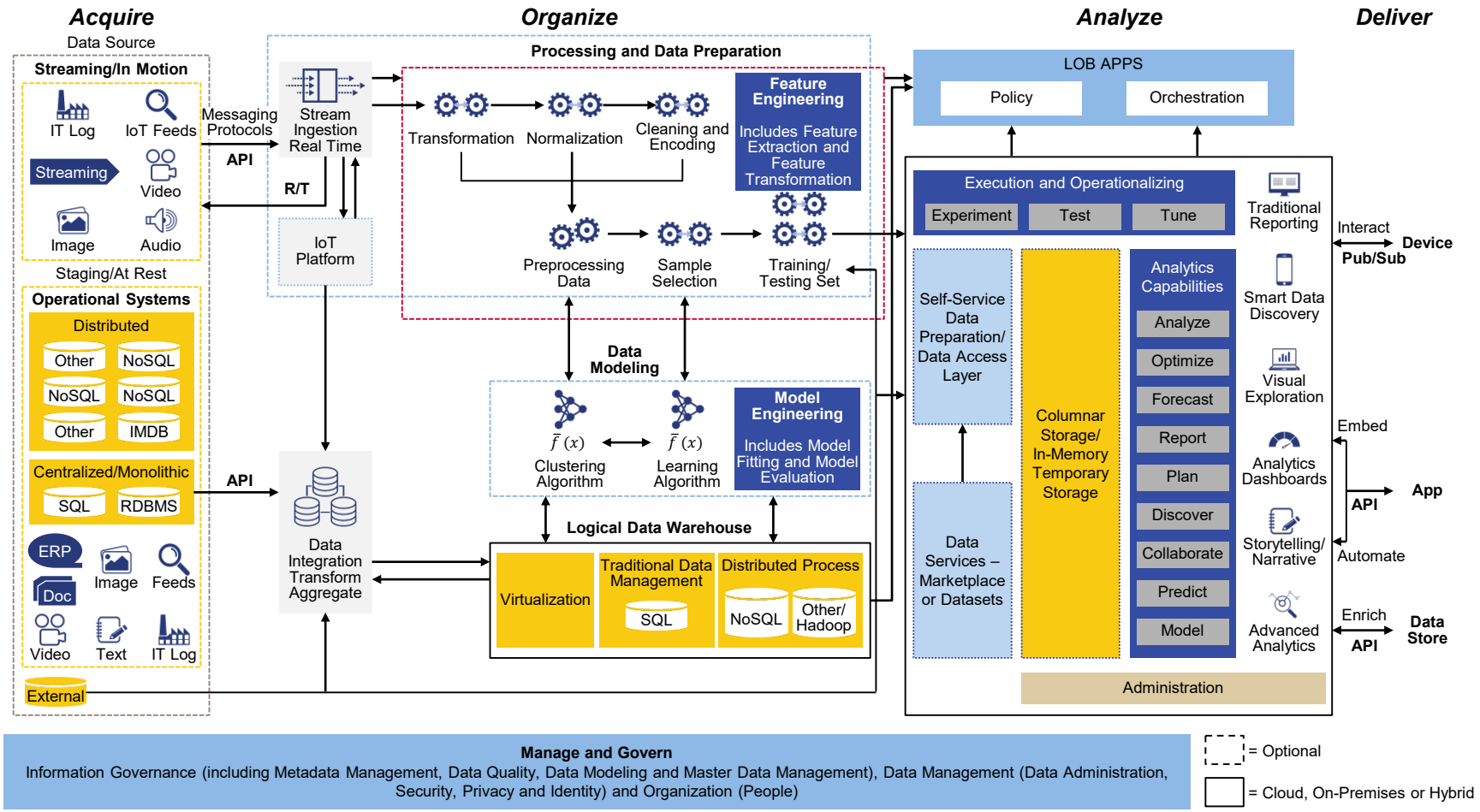
Amazon Retail FA**C**E

Six-page brief



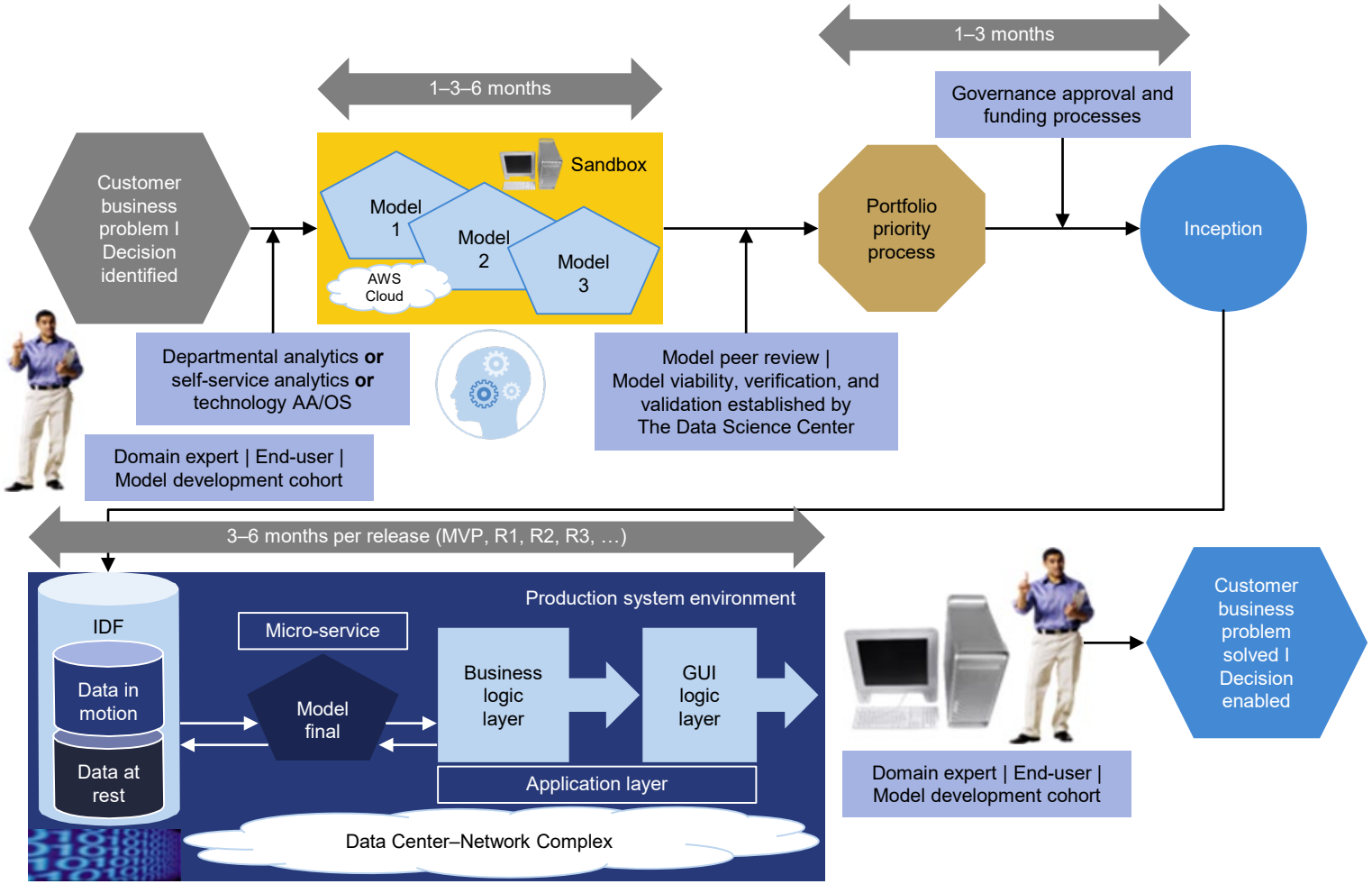
MS PPT

Amazon Retail FACE



Amazon Retail FACE

Analytics project (delivery execution) management



DataScience@SMU