

The background of the image is a blurred laptop keyboard. Overlaid on the keyboard is a stylized globe with a network diagram. The globe is rendered in a painterly style with yellow, brown, and blue tones. A network of white dots connected by thin white lines is superimposed over the globe and extends across the upper portion of the image. The word "CODIFYD" is written in a bold, dark blue, sans-serif font across the middle of the image, partially overlapping the globe and the network lines.

CODIFYD

Driving Business Impact by
Adopting Advanced Analytics and
Machine Learning



Santosh Abraham

Analytics Lead

A Bit About Me

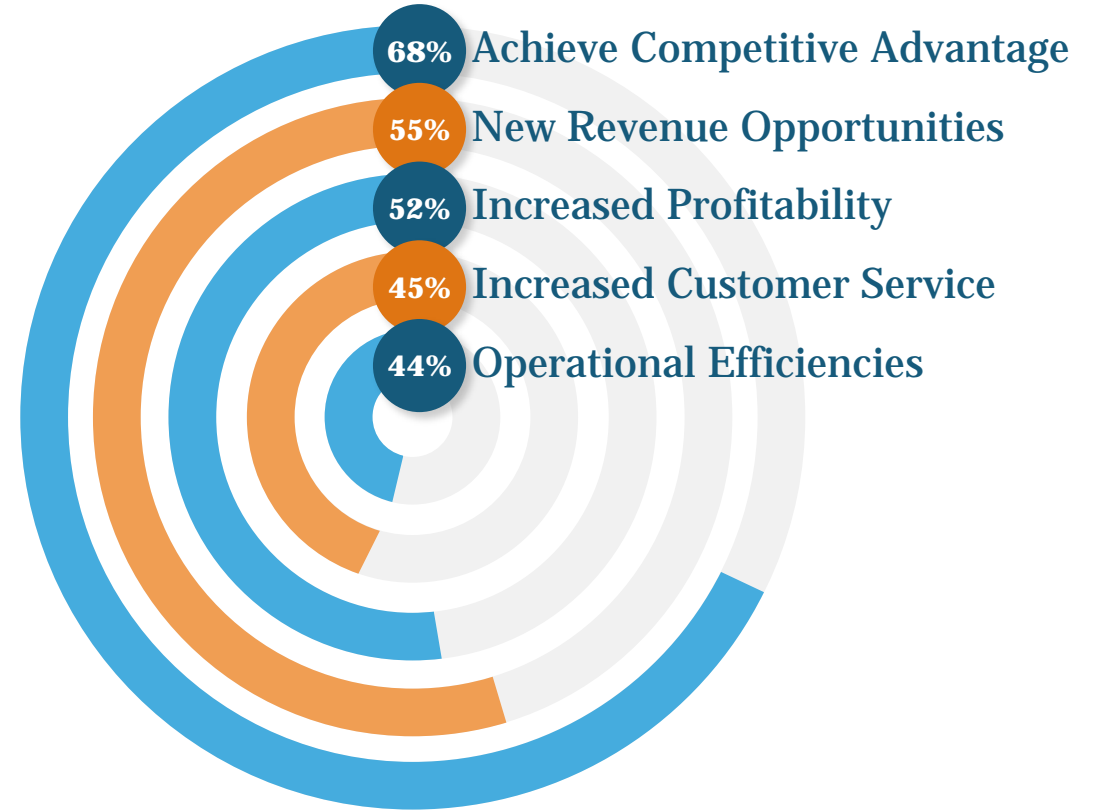
Part 1 : Quick Dive – Analytics and Machine Learning

Why Analytics?

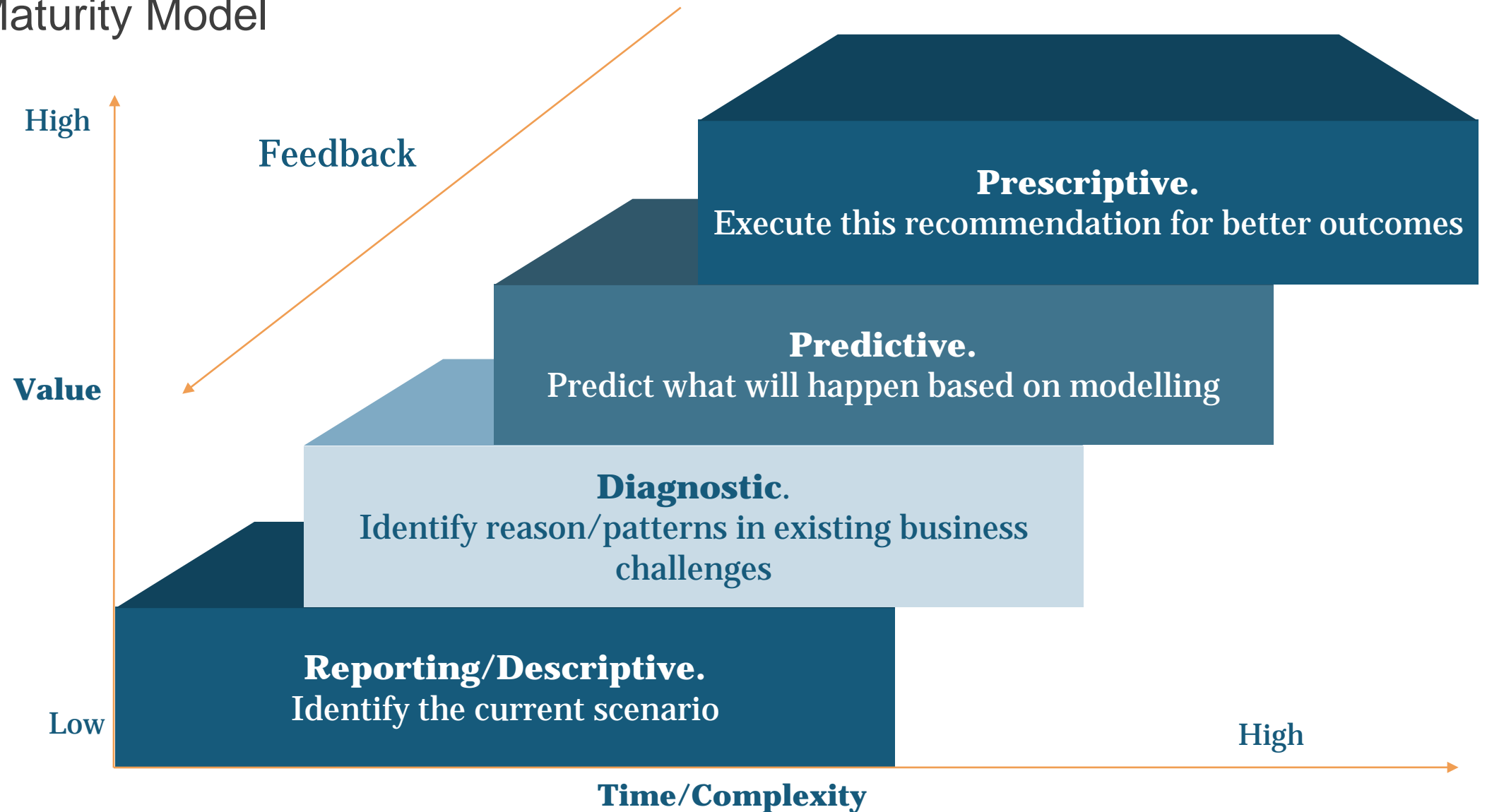
Gartner

“By 2020, information will be used to reinvent, digitalize or eliminate 80% of business processes and products from a decade earlier. By 2017, more than 30% of enterprise access to broadly based data will be delivered by insight companies, serving context to business decisions.”

Benefits of Analytics

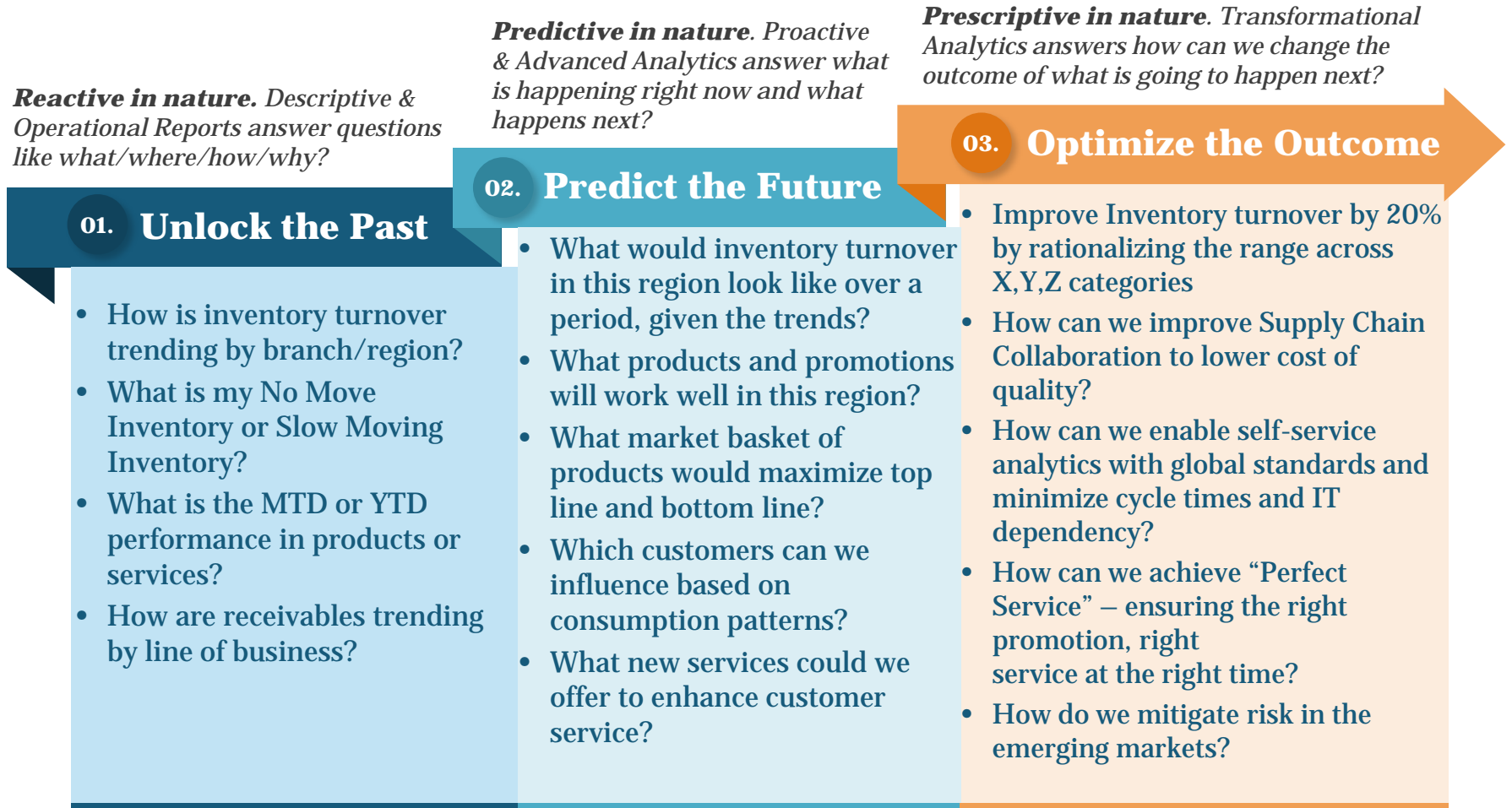


Analytics Maturity Model



3 Stages of Organizational Analytics Maturity

**HELPS
ORGANIZATIONS
TRANSFORM**



**OK, then, what is
*Machine Learning?***

A thin vertical orange line is positioned on the right side of the image, extending from the bottom towards the middle.

Components of ML



Algorithm + Experiences (training data)

- Regression, K-means clustering, support vector machines
- Neural networks
- Deep learning networks
- Data – Volume, Velocity and Variety

Generalized strategy for learning

How to :

- Classify
- Auto Generate
- Predict

Identify patterns, associations and insights + Develop new model based on these inputs

- Identify objects in an image
- Recognizing text/audio
- Interpolating missing data
- Making predictions



Applying ML to a Business Problem

Let's take the Return on Marketing Investment (ROMI) problem

What is the Problem?



“Half the money I spend on advertising/marketing is wasted...the problem is I don’t know which half.”

Paid Media

- Display
- Social Ads
- Video
- In App Ads
- Search
- Retargeting
- CSEs
- Affiliates
- Email

Owned Media

- Referrals
- Organic Search
- Direct Nav
- Organic Social
- Content

Audience and CRM Data

Transaction & Sales Channels

- Websites
- Call-Center
- Kiosk
- Brick-n-Mortar

SOLUTION FRAMEWORK ANALYTICS ARCHITECTURE



Algorithmic Models
Visualization
Data Manipulation



Attribution
Reporting



Cross Channel
Optimization



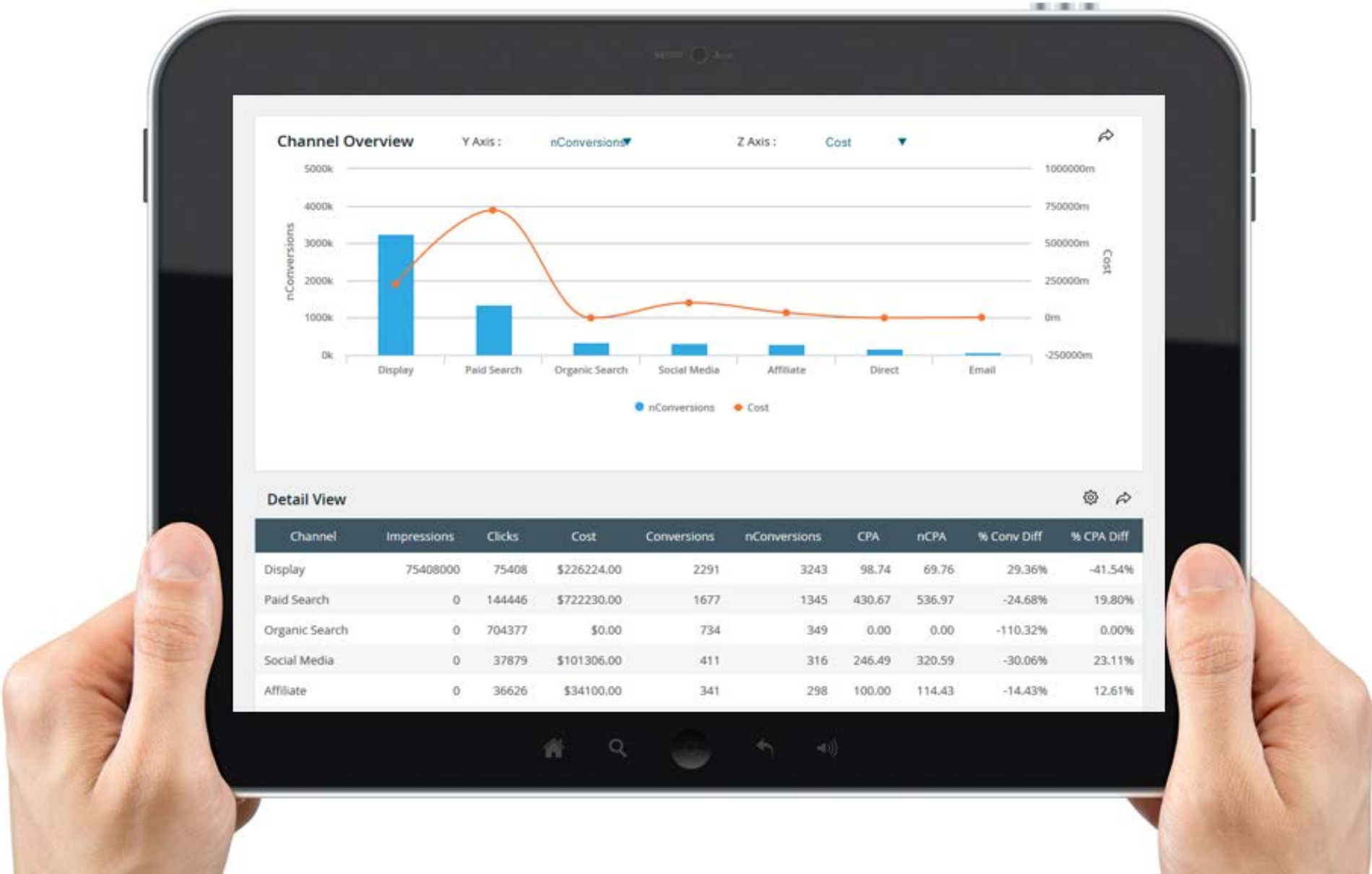
Audience
Insights

Data Store

Analytics Engine

Reporting

ANALYTICS
SOLUTION

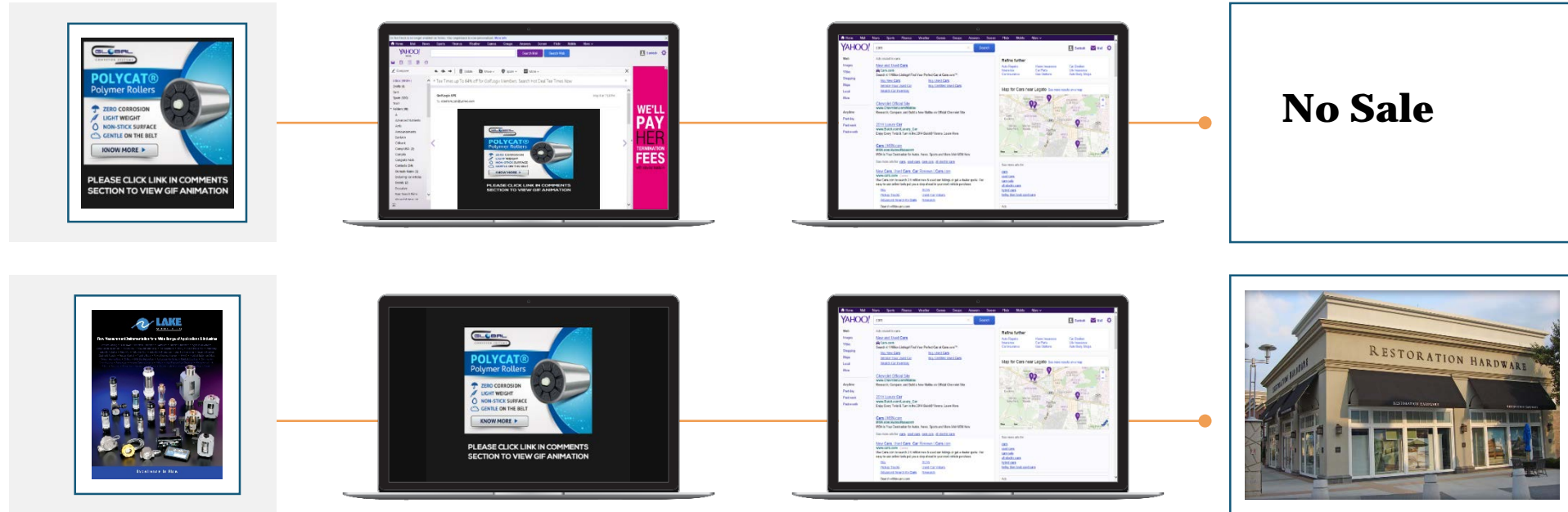




Problem solved, right?

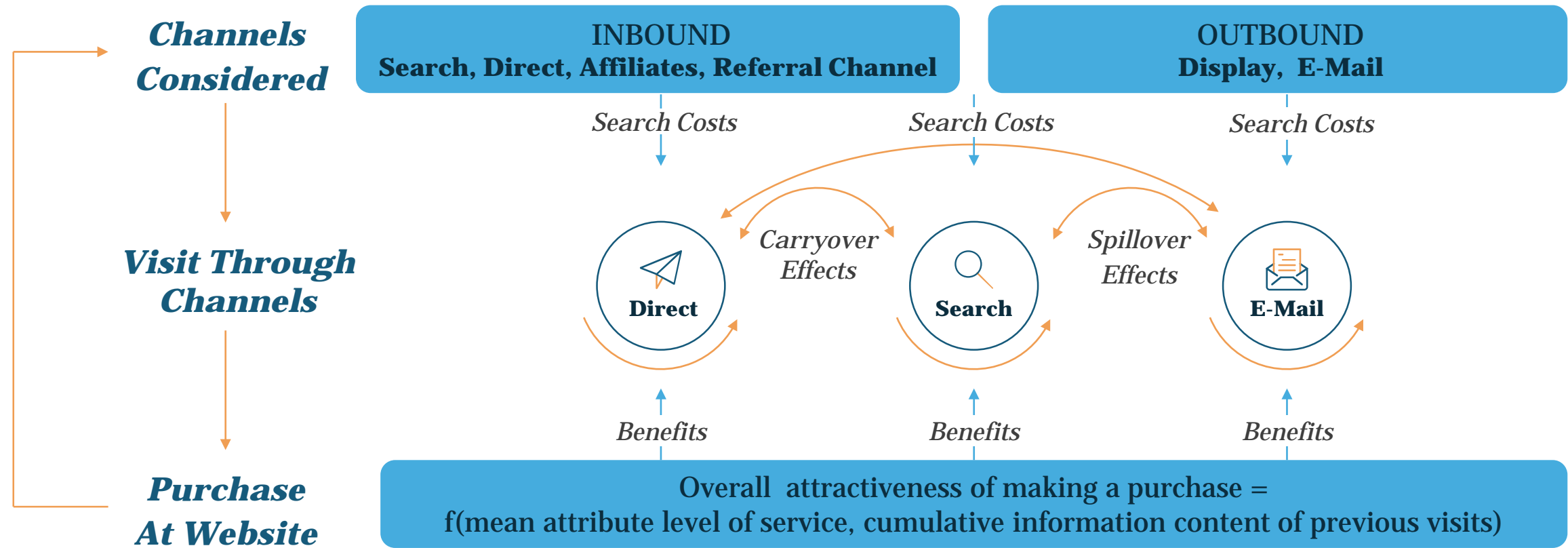
...not completely.

Let's *leverage* Machine Learning.



Training Model

Strategy for learning





Algorithm

$$P(B_{ijn} = 1, V_{in} = j | C_i) = \Pr(B_{ijn} = 1 | C_i, V_{in} = j) P(V_{in} = j | C_i)$$

$$L(B | \theta) = \prod_{n=1}^{N_i} \prod_{i=1}^I \prod_{j=i}^J \sum_{k=1}^{2^J - 1} P(C_i = H_k | \alpha, \Sigma) \times \left[b_{ijn1}^{B_{ijn}} b_{ijn0}^{(1-B_{ijn})} \right]$$

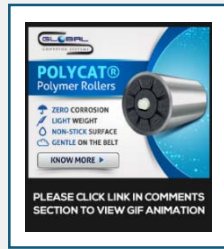
where

$$b_{ijn1} = P(V_{in} = j | C_i; \beta, \mu, \tau) P(B_{ijn} = 1 | C_i, V_{in} = j; \gamma, \lambda)$$

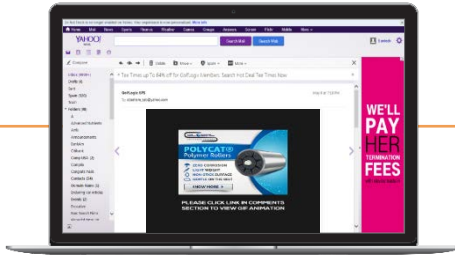
$$b_{ijn0} = P(V_{in} = j | C_i; \beta, \mu, \tau) [1 - P(B_{ijn} = 1 | C_i, V_{in} = j; \gamma, \lambda)]$$

Machine Learning – Pattern recognition

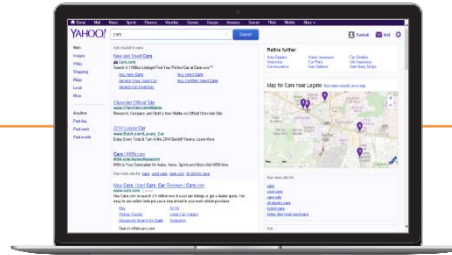
Recognizing new patterns



C1F1CR1



EC1G2M1



PC1KG1CR5



C1F1CR1EC1G2M1PC1KG1CR5



Display(D)

- Campaign(C1)
- Site(S1)
- Placement(P1)
- Format(F1)
- Creative(CR1)



Email(E)

- Campaign(C1)
- Segment(G2)
- Primary Message(M1)
- Secondary Message(SM1)



Paid Search(P)

- Campaign(C1)
- Provider(PR2)
- Keyword Group(KG1)
- Keyword(K1)
- Creative(CR5)

Hyper Personalization: The Power of Recommendations



- Deliver personalized content
- 80% of Content hours streamed driven by personalization engine
- Customer retention and Engagement estimated to be worth \$1 Billion

Part 2 : Successful Analytics and Data-Led Transformation

Is Data the new oil?

Let's look at ROI of Analytics based on McKinsey study.

“Data, whether big or small, is a raw material; it does not have intrinsic value. To extract its value, data must be refined through analysis.”

President – Multi-Billion dollar Distributor

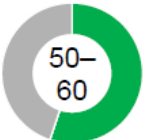
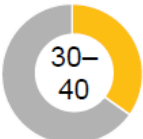
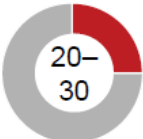
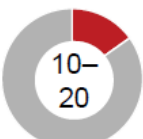
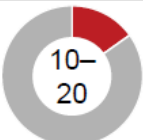
“At **Company X** we hire many ‘analysts’. We have analysts in every department. It's really unbelievable. What they are doing and what I think analysts should be doing are two different things.”

“What we have are excellent data pullers. Our analysts can pull data from our systems all day long. However, interpreting the figures and giving us ‘Top 3’ suggestions to make **Company X** better...that is another story.”

“If you truly understand the data, you should be chock-full of ideas and plans to make us better rather than just reciting data we can easily read.”



There has been uneven progress in capturing value from data and analytics

	Potential impact: 2011 research	Value captured %	Major barriers
Location-based data	<ul style="list-style-type: none"> ▪ \$100 billion+ revenues for service providers ▪ Up to \$700 billion value to end users 		<ul style="list-style-type: none"> ▪ Penetration of GPS-enabled smartphones globally
US retail¹	<ul style="list-style-type: none"> ▪ 60%+ increase in net margin ▪ 0.5–1.0% annual productivity growth 		<ul style="list-style-type: none"> ▪ Lack of analytical talent ▪ Siloed data within companies
Manufacturing²	<ul style="list-style-type: none"> ▪ Up to 50% lower product development cost ▪ Up to 25% lower operating cost ▪ Up to 30% gross margin increase 		<ul style="list-style-type: none"> ▪ Siloed data in legacy IT systems ▪ Leadership skeptical of impact
EU public sector³	<ul style="list-style-type: none"> ▪ ~€250 billion value per year ▪ ~0.5% annual productivity growth 		<ul style="list-style-type: none"> ▪ Lack of analytical talent ▪ Siloed data within different agencies
US health care	<ul style="list-style-type: none"> ▪ \$300 billion value per year ▪ ~0.7% annual productivity growth 		<ul style="list-style-type: none"> ▪ Need to demonstrate clinical utility to gain acceptance ▪ Interoperability and data sharing

¹ Similar observations hold true for the EU retail sector.

² Manufacturing levers divided by functional application.

³ Similar observations hold true for other high-income country governments.

SOURCE: Expert interviews; McKinsey Global Institute analysis



AI/ML are Buzzworthy terms and everybody is trying to jump on the bandwagon.



Companies are making the mistake of layering in AI without a strong basis in analytics and data.



“Companies with **strong basic analytics** - make breakthroughs in complex and critical areas after layering in AI/ML.”

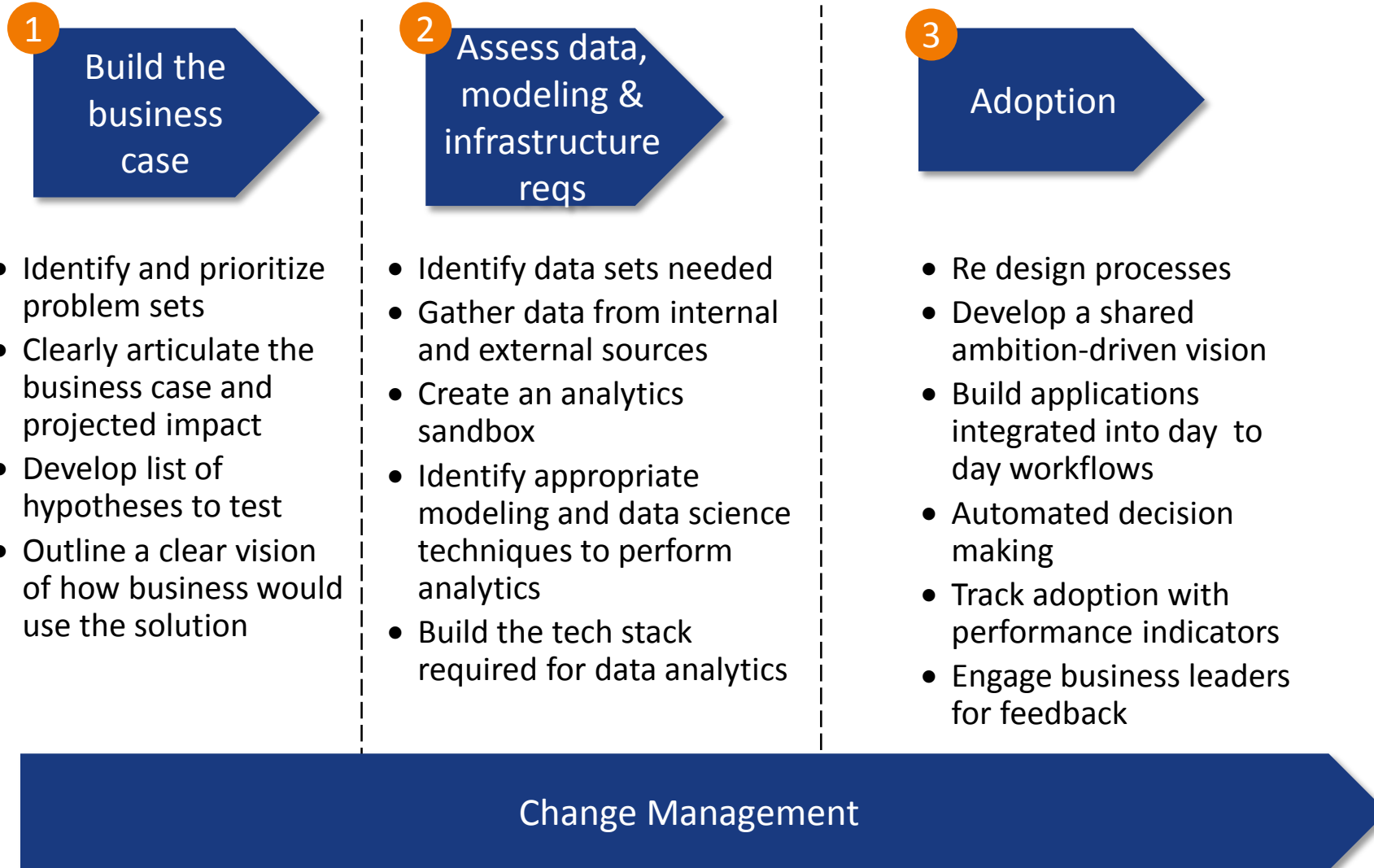
-Harvard Business Review Study



A Framework for Successful Data and Analytics-Based Transformation

Analytics 360 + CoE

Model for successful data and analytics transformation



Model for successful data and analytics transformation

1

Build the business case

- Identify and prioritize problem sets
 - Clearly articulate the business case and projected impact
 - Develop list of hypotheses to test
 - Outline a clear vision of how business would use the solution
- Do you know how Analytics can disrupt your industry/existing business model?
 - What can be learned from other industries that are further along?
 - Market scan of your competitors and data/analytics vendors
 - Potential for players to upend prevailing business models
 - Where and how can analytics create value?
 - Do you have clarity on what business value you are trying to create?
 - Do you have identified business cases that outline this?
 - Have you sized and prioritized these use cases?

Model for successful data and analytics transformation

1

Build the business case

- Identify and prioritize **problem sets**
- Clearly articulate the business case and projected impact
- Develop list of hypotheses to test
- Outline a clear vision of how business would use the solution

Pricing Intelligence

Pricing Analytics (Am I selling my products at right prices?)

- Evaluate historical data for pricing trends
- Plan product pricing
- Revenue maximization by identifying pricing leaks
- **Dynamic pricing**
- **Promotion Pricing Analytics**
- **Pricing Elasticity / Sensitivity**

Models

ARIMA, SARIMAX, Dickey-Fuller, Box-Jenkins, Spectral Analysis, Holts Winter

Logistics Intelligence

Transportation Analytics (How to optimize transportation routes and loads)

- Optimizing transport routes including backhaul
- Optimizing shipment schedules.
- Optimizing modes of logistics

Procurement Intelligence

Procurement Analytics (How to achieve lowest landed cost and secure long-term high quality supplier partners)

- Scoring models for vendor quality, cost, and stability

Inventory Intelligence

Finished Inventory Optimization (What stock should I hold and where should I position it?)

- Inventory budget optimization
- Safety stock level recommendations
- Segment inventory for tailored and customized fulfillment strategies by customer type

Replenishment Planning Analytics (What, when, and where should I ship?)

- Integrated planning at the retailer, distributor, and channel level
- Optimize fulfillment logistics to account for handling, storage or warehouse constraints

2

Assess data,
modeling and
infrastructure
requirements

- Identify data sets needed
 - Gather data from internal and external sources
 - Create an analytics sandbox
 - Identify appropriate modeling and data science techniques to perform analytics
 - Build the tech stack required for data analytics
- Have you been able to get the right data and get it ready for analysis?
 - Have you been able to agree on “a single source of truth”?
 - What kind of data is required to improve competitive advantage?
 - Do you have unique identifiers to link data sets?
 - Transac data and customer profiles
 - Do you have data governance and a process to ensure the rightful owners have direct access?
 - Have you created a data lake? How did you go about doing it? How to prevent boiling the ocean?

2

Assess data,
modeling and
infrastructure
requirements

- **Identify data sets needed**
- **Gather data from internal and external sources**
- Create an analytics sandbox
- **Identify appropriate modeling and data science techniques to perform analytics**
- Build the tech stack required for data analytics

Demand Intelligence

Demand Analytics
(How is my forecast tracking with actual sales?)

Models

ARIMA, SARIMAX, Dickey-Fuller,
Box-Jenkins, Spectral Analysis,
Holts Winter

Data Required

Transaction Data
Master Data - Stores
Master Data - Distributors
Master Data - Manufacturers
Demographics Data
Holidays Data

Procurement Intelligence

Procurement Analytics
(How to achieve lowest landed cost and secure long-term high quality supplier partners)

Models

Regression – Linear, Decision Tree,
Random Forest
Classification – SVM, Naïve Bayes,
Logistics Regression

Data Required

AP Data
Sourcing Data
Credit Data
Risk Data
HSE Data

Logistics Intelligence

Transportation Analytics
(How to Optimize transportation routes and loads)

Models

Travelling Salesman, Vehicle
Routing, Route inspection, Vehicle
rescheduling

Data Required

Fleet Management System Data
Route Planning System Data
Order / Delivery Data
Inventory Data
Warehouse Data

2

Assess data,
modeling and
infrastructure
requirements

- Identify data sets needed
- Gather data from internal and external sources
- Create an analytics sandbox
- Identify appropriate modeling and data science techniques to perform analytics
- **Build the tech stack required for data analytics**

Cloud



Big Data



Analytical Tools



Visualization





Databases



Adoption

- Re design processes
- Develop a shared ambition-driven vision
- Build applications integrated into day to day workflows
- Automated decision making
- Track adoption with performance indicators
- Engage business leaders for feedback

How do I get started? Examples we have seen work.

		What	Outcomes	Time
01.		ANALYTICS 360 is a discovery engagement that helps organizations identify problem sets that can be solved through analytics, prioritize them based on business impact and help lay the roadmap for solution building and implementation	<ul style="list-style-type: none"> • Identified problem sets that can be solved by applying analytics • Prioritization based on business impact • High level solution architecture • Identification of data sources and data transformation required • Tool and infrastructure recommendations 	6 to 8 weeks
02.		Center of Excellence (CoE) builds on the 360 and puts together the team, technology and governance required to execute on the roadmap	<ul style="list-style-type: none"> • CoE cross-functional starter team • 3 Year vision and 6 month action plan • Analytics and data frameworks • Initial Analytics POC • Solution build and deployment • Business user training and operationalizing solution for ongoing use. 	6 months -1 year

Part 3: Case Study

Case – Product recommendation engine

Business – Large Electrical B2B distributor

Challenge

Lost sales due to product range rationalization
Reduced SKUs by 50,000
Sales team did not know product alternates existed

Data Challenges

Product data in various systems
Missing attributes, inaccurate data

STRATEGY

- Data enrichment, transformation and integration
- Leverage ML based matching algorithms to create product alternates

MODELLING

- Category Prioritization
- Weights and rules for data matching
- Data gaps resolved
- Matching algorithms using ML
- Validation by category managers.

BUILD

- Integrate output into recommendation engine
- Product alternates matching recommendation with 5 levels.

OUTCOME

- Over 80% improvement in product coverage
- Average item matching improved over 300%
- Improved product attribution fill ratios

Case – E-commerce sales conversion

Business – Online B2B office supplies provider

Challenge

- Site traffic increased
- Revenue was stagnant

Business Problem

- To discover the issue causing this stagnation in conversion
- There was a problem with conversion funnel. A thorough analysis of user behavior throughout the website was needed

Data considered

- Historical data related to conversion funnel
- Behavioral Flow data

ANALYSIS

We found that users who knew what they were looking for converted at a higher rate, however when searching using website's search function, they exited more often than moving on to product page. Exit rate for search query was 91%.

RECOMMEND

Optimize search functionality

BUILD

Built multi faceted search
Mapped nodes to search query
Re built TAXONOMY

OUTCOME

Month on month revenue increase by 17%
Conversion rate increased by 10%

Examples in Wholesale Distribution

Making sales efficient

Sales calls can be more productive. The average sales visit is 20 minutes, so a wholesale distribution sales rep may only have time to introduce three to five of the thousands of products in their catalog. Predictive analytics provides the sales rep with the three to five products that the customer is most apt to purchase.

Special Promotions

Special promotions are being used more frequently in wholesale distribution. When introducing new categories, sales reps can leverage data about the optimal customer profiles to introduce new products and penetrate the market.

Predict future sales

Distributors can leverage Big Data to analyze which customers purchased certain products to predict future sales. This is a strong service for suppliers, who typically only have data on quantities sold, and enables them to take advantage of precision marketing and customer segmentation to improve product strategy and grow their business.



CODIFYD

Thank You!

codifyd.com 303 E. Wacker Suite 950 Chicago, IL 60601