THE ART & SCIENCE OF PREDICTIVE ANALYTICS AND DATA SCIENCE

26 March 2019, Singapore

Dr Dan Putler, Chief Data Scientist, Alteryx
Vincent Toh, Team Lead, Sales Engineering, Alteryx

alteryx | The Thrill of Solving

9:00 AM	Welcome & Introduction
9:10 AM	Developing an In-House Predictive Capability
10:00 AM	Addressing the Data Problem
10:30 AM	Break
10:45 AM	Real-World Example Beyond Basic BI
11:30 AM	Putting the Power of Predictive into Production
11:45 AM	Q/A
12:00 PM	End of Event

DEVELOPING AN IN-HOUSE PREDICTIVE ANALYTICS CAPABILITY

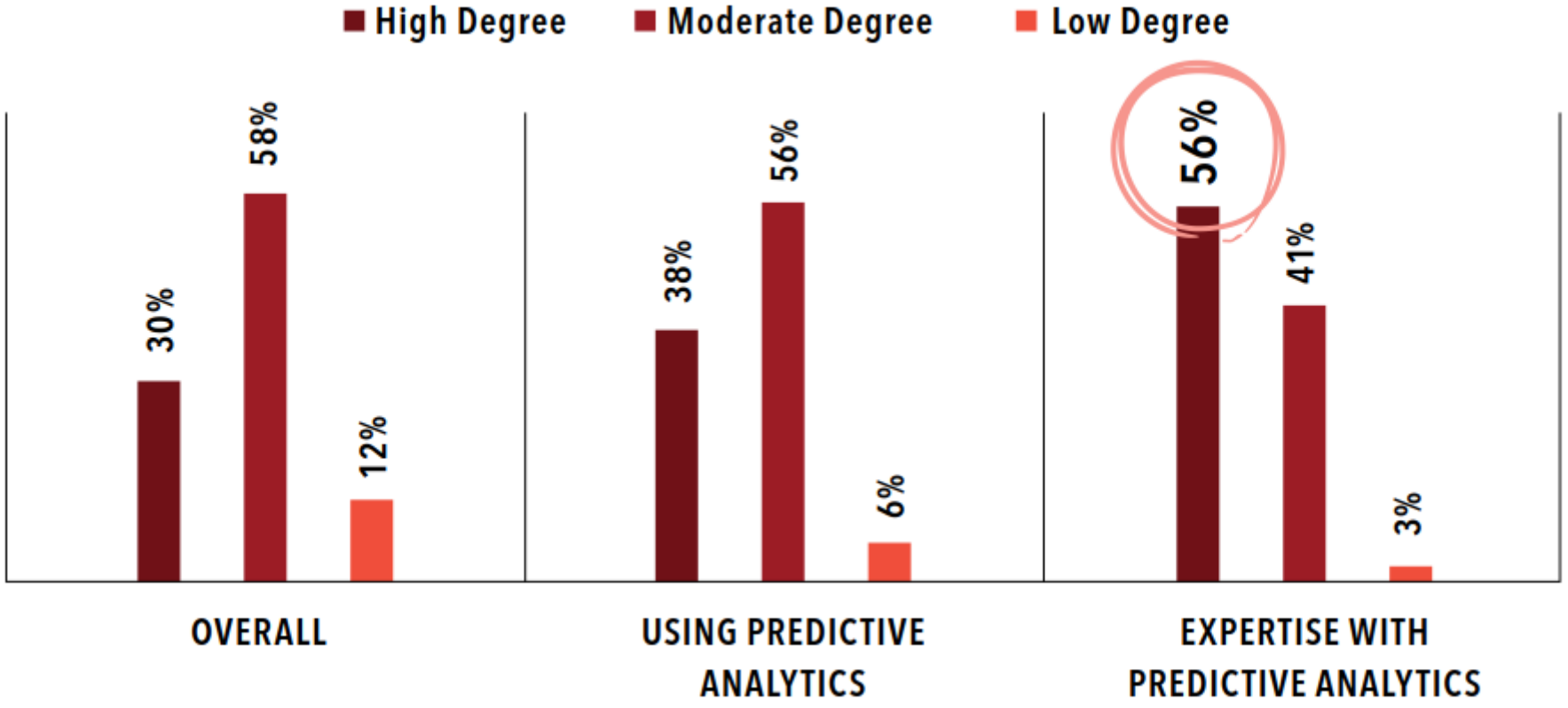
alteryx | The Thrill of Solving

- What is predictive analytics and how does it differ from BI?
- The benefits and difficulties of creating a predictive analytics capability
- Four steps to build in-house predictive analytics expertise

BI AND REPORTING VS. PREDICTIVE ANALYTICS

- Same underlying goal, to inform management decision making
  - Requires understanding of business issues in order to provide the right information to inform business decisions
- What the two approaches provide is different
  - BI and reporting provide information that is currently available
  - Predictive analytics provides information that is not currently available, but which can be reasonably predicted

PREDICTIVE ANALYTICS INFLUENCE ON PROGRAM SUCCESS



Source:
TDWI 2017

PREDICTIVE ANALYTICS - OPPORTUNITY AND CHALLENGE

Return on Investment

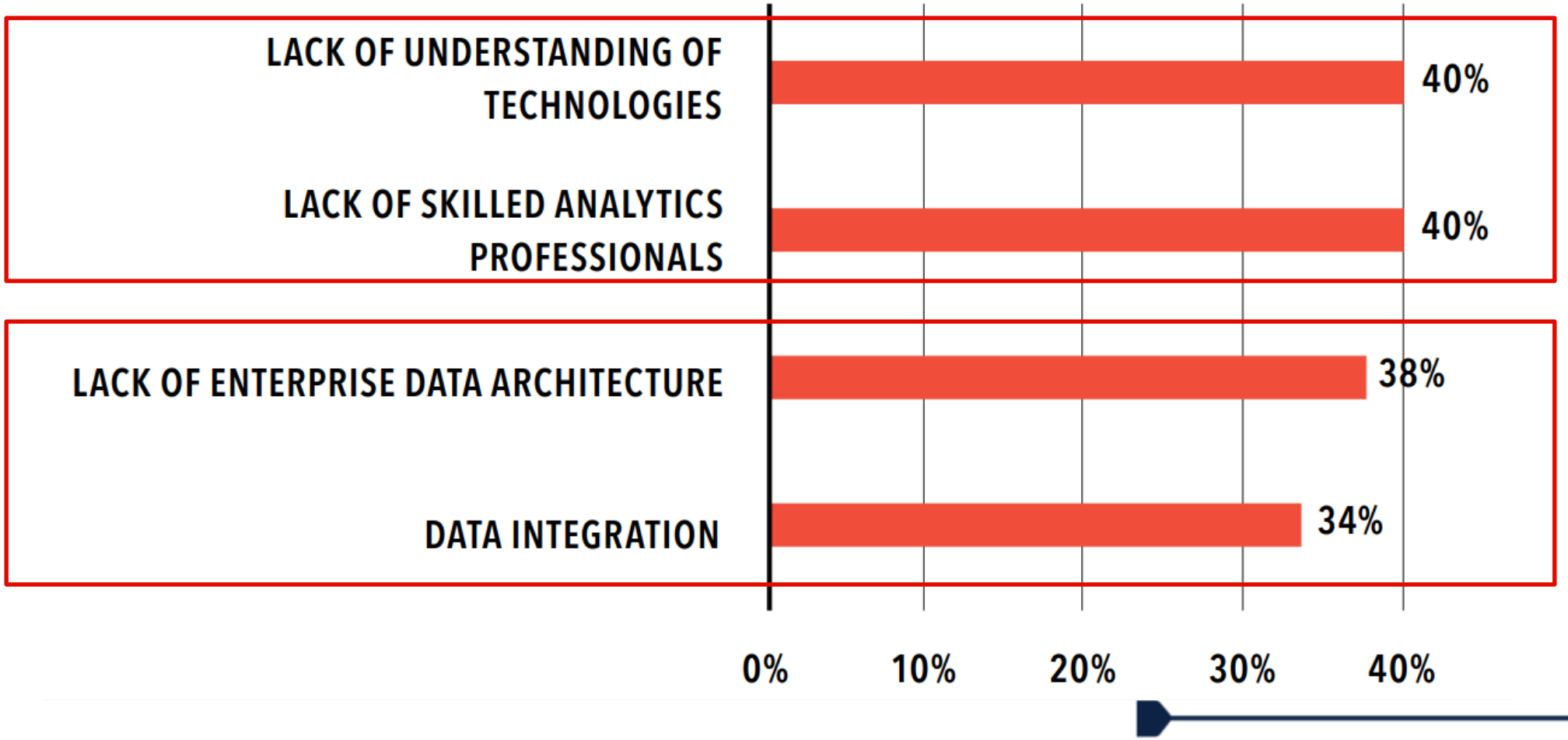
250%

2.5X

35%

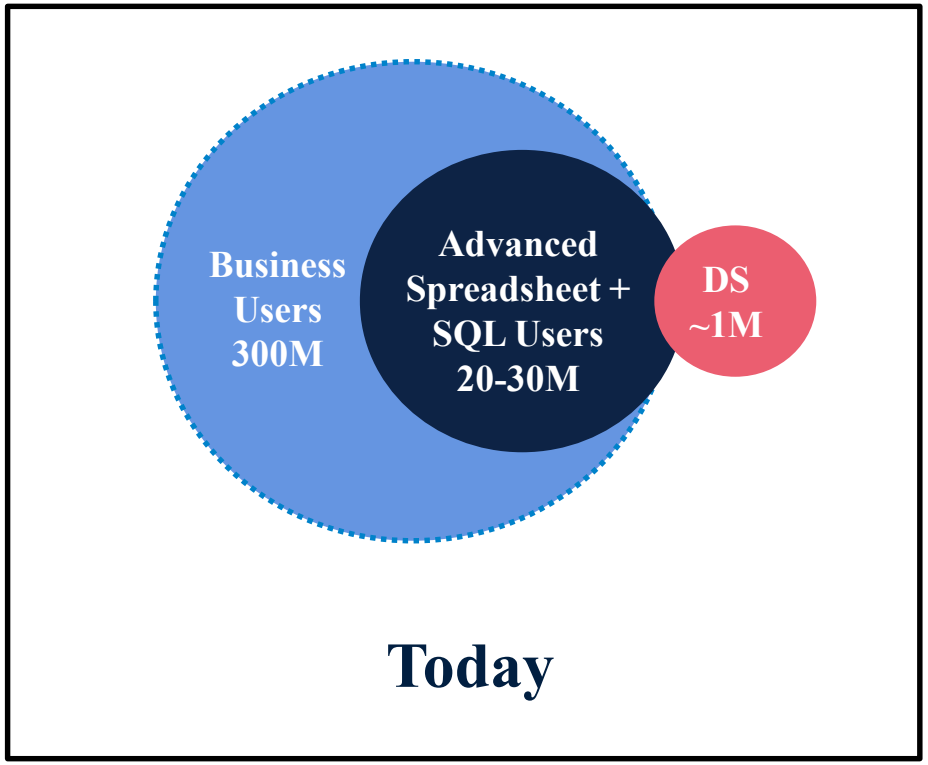
Implementing Predictive Analytics

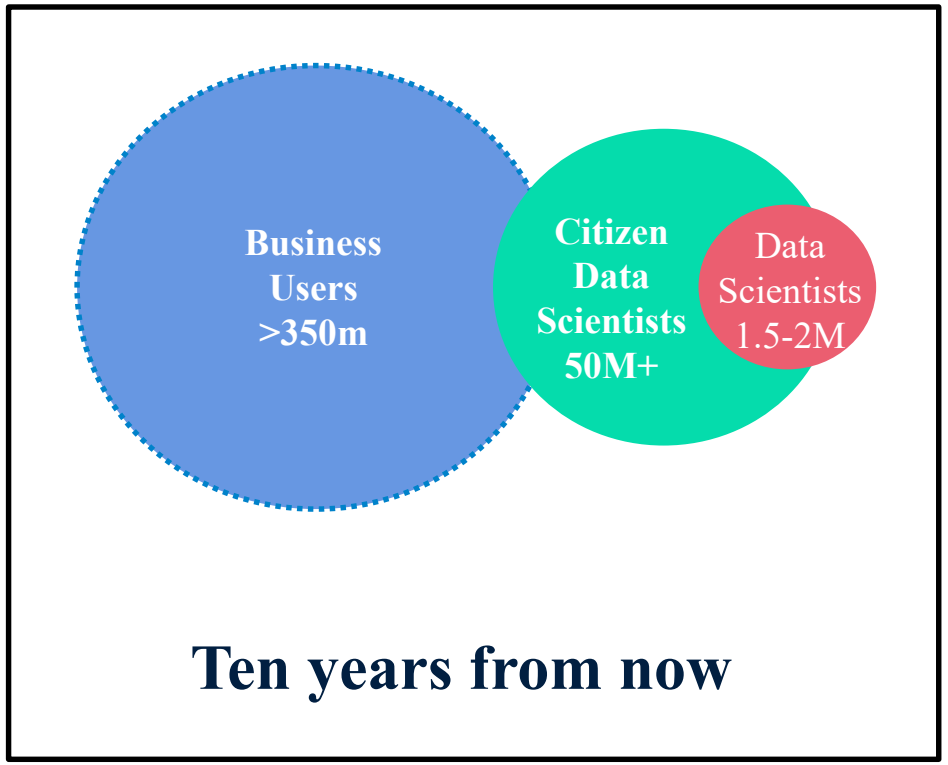
FACTORS SLOWING PREDICTIVE ANALYTICS ADOPTION



Source:
TDWI 2017

THE SCARCITY OF DATA SCIENTISTS AND WHAT TO DO





FOUR STEPS TO DEVELOP PREDICTIVE EXPERTISE

1. Start small and take a “learning by doing” approach
2. Develop an initial list of possible predictive analytics projects that address frequent and important business decisions in your organization
3. Select projects from the initial list that can make use of well-known metrics
4. Compare the results of a new predictive analytics based process to the old process used to make the decision

1. START SMALL AND TAKE A "LEARNING BY DOING" APPROACH

- Start with one to four people involved
  - Have the belief that predictive analytics could improve their organizations, be willing to acquire some additional skills, and have enough time to experiment with the methods
  - Understanding the business and data is more important
    - » Some upskilling on technical capabilities may be needed
- Only a small number (one to a handful) of projects are undertaken that are often "skunkworks" in nature
  - Minimizes the initial investment, prevents over inflated expectations, allows the organization to begin to develop expertise in the area they can leverage in the future, and gain a realistic sense of the potential returns to predictive analytics

2. DEVELOP A LIST OF POSSIBLE PREDICTIVE ANALYTICS PROJECTS

- Thinking through how "predicted information" can be used to help decision making is a crucial first step to developing a successful in-house predictive analytics program
- It is often useful to start the process by examining your organization's key performance indicators, or KPIs
  - Think through what underlying business decisions drive a particular KPI, and then think of ways predicted information could better inform those decisions
- Example: Personal loans issued by a credit union

3. SELECT PROJECTS THAT MAKE USE OF WELL-KNOWN "METRICS"

- A predictive model is only as good as the data. Selecting the right data is crucial
- Determining the appropriate metrics can be challenging, but there is often no need to re-invent the wheel
- Many verticals and/or domains have "go to" metrics for predicting information
  - Recency, frequency, and monetary value (or RFM) in marketing
  - Credit scores in financial services
  - Elo ratings in sports

4. COMPARE RESULTS OF THE NEW PROCESS WITH THE OLD PROCESS

- Testing and experimentation are an essential part of predictive analytics
- There are two basic testing approaches, A/B testing and retrospective testing
  - A/B testing involves the use of two (or more) different samples with one sample being the "treatment" group where a predictive model is employed for decision making, and the other a "control" group that uses the current decision making process
  - Retrospective testing uses historical data from two time periods, where models are created in the earlier period, and a comparison is made between the decisions actually made, and those that we would make with the predicted information

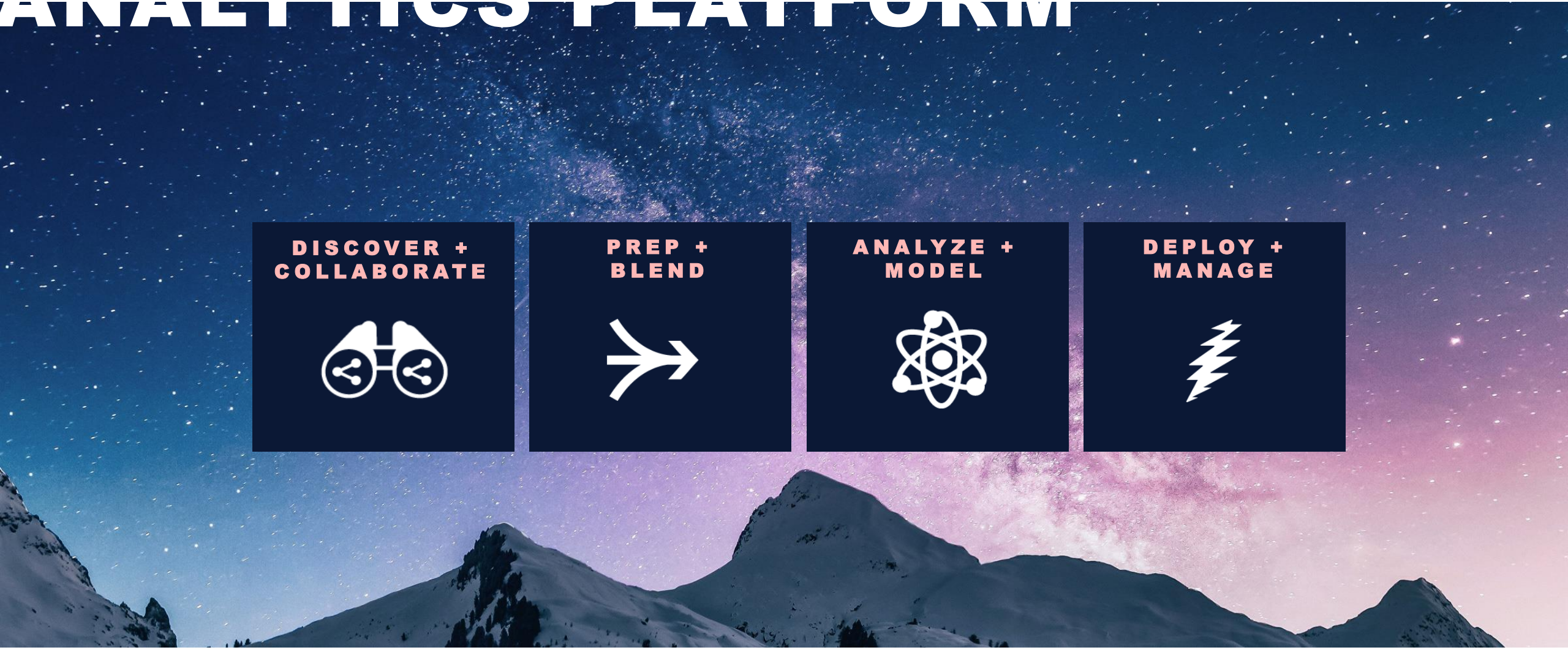
KEY TAKE-AWAYS

- Predictive analytics and traditional BI and reporting have the same objective, improve decision making by better informing decisions, but predictive analytics can provide information that BI and reporting cannot
- Predictive analytics can have significant returns, but there are challenges that need to be overcome
- Taking a four step approach to developing expertise in predictive analytics can help overcome some of those challenges

ADDRESSING THE DATA PROBLEM

alteryx | The Thrill of Solving

FLEXIBLE END-TO-END ANALYTICS PLATFORM



alteryx | The Thrill of Solving

DISCOVER, CATALOG AND SHARE

YOUR DATA

DISCOVER

Disparate Data / Apps

CATALOG

Lineage / Governance

COLLABORATE

Crowdsource / Visibility

ALTERYX CONNECT

SOCIAL, SMART DATA DISCOVERY

YOUR DATA SOURCES

salesforce
amazon web services
ORACLE
tableau SOFTWARE
Microsoft SQL Server
snowflake
alteryx ANALYTICS GALLERY

CURATED TRIBAL KNOWLEDGE…

Assemble → **DISCOVER** → **ACCESS** → **TRUST** → **SHARE**

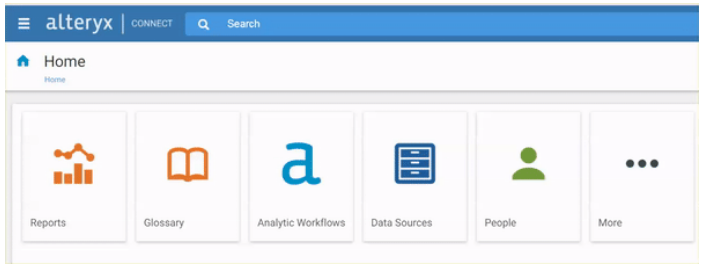
SEARCH

"WHERE DO I FIND THE DATA I NEED FOR MY ANALYSIS"?

"WHAT ARE THE RELATIONSHIPS BETWEEN MY DATA SOURCES AND WHAT DOES IT MEAN?"

"HOW CAN I TRUST THIS DATA AND ENSURE THAT IT IS RELEVANT AND ACCURATE?"

"HOW CAN I COLLABORATE AND SHARE MY ANALYSIS AND FURTHER ENRICH IT?"



alteryx | The Thrill of Solving

ALTER YOUR ANALYTICS LIFECYCLE

BEFORE

IT

- Data Sources
- Data Cleansing Scripts and Tools
- Prep and Blending Scripts and Tools

Informatica, Talend, Paxata, Trifacta

GIS Specialists

Spatial Tools

Esri, Pitney Bowes, SAP

Data Scientists

Predictive and Prescriptive Scripts and Models

SPSS, SAS, R, Python

IT and DevOps

Production and Output Development

Java, PHP, C++, Ruby

Business Analysts

Visualizations & Reports

Microsoft, Tableau, Qlik

ALTERYX SELF-SERVICE

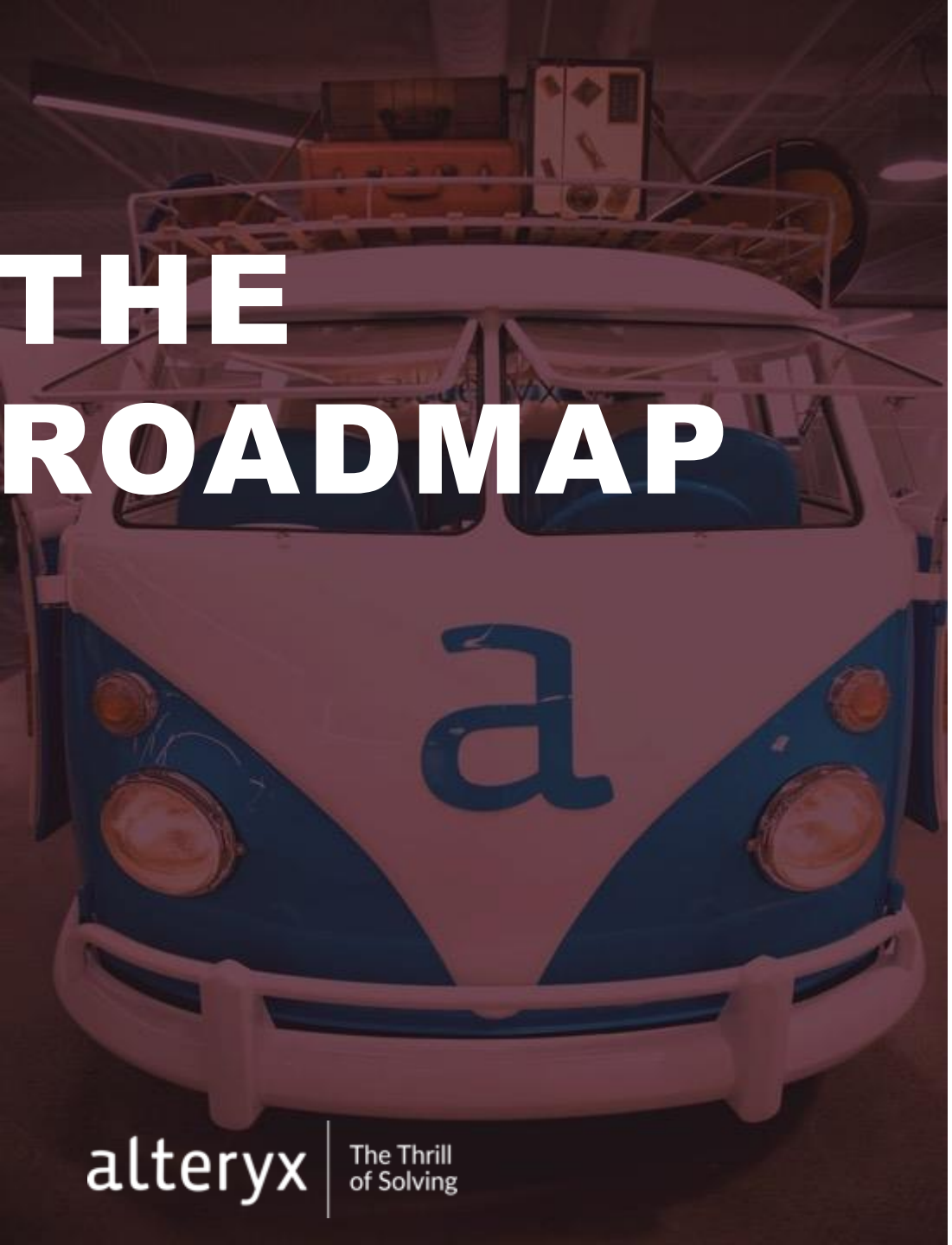
Analyst

Find the right data asset → Data Cleansing → Prep and Blending → Spatial Data → Predictive and Prescriptive → Production and Output

NEW INSIGHTS
BETTER DECISIONS
MORE VALUE

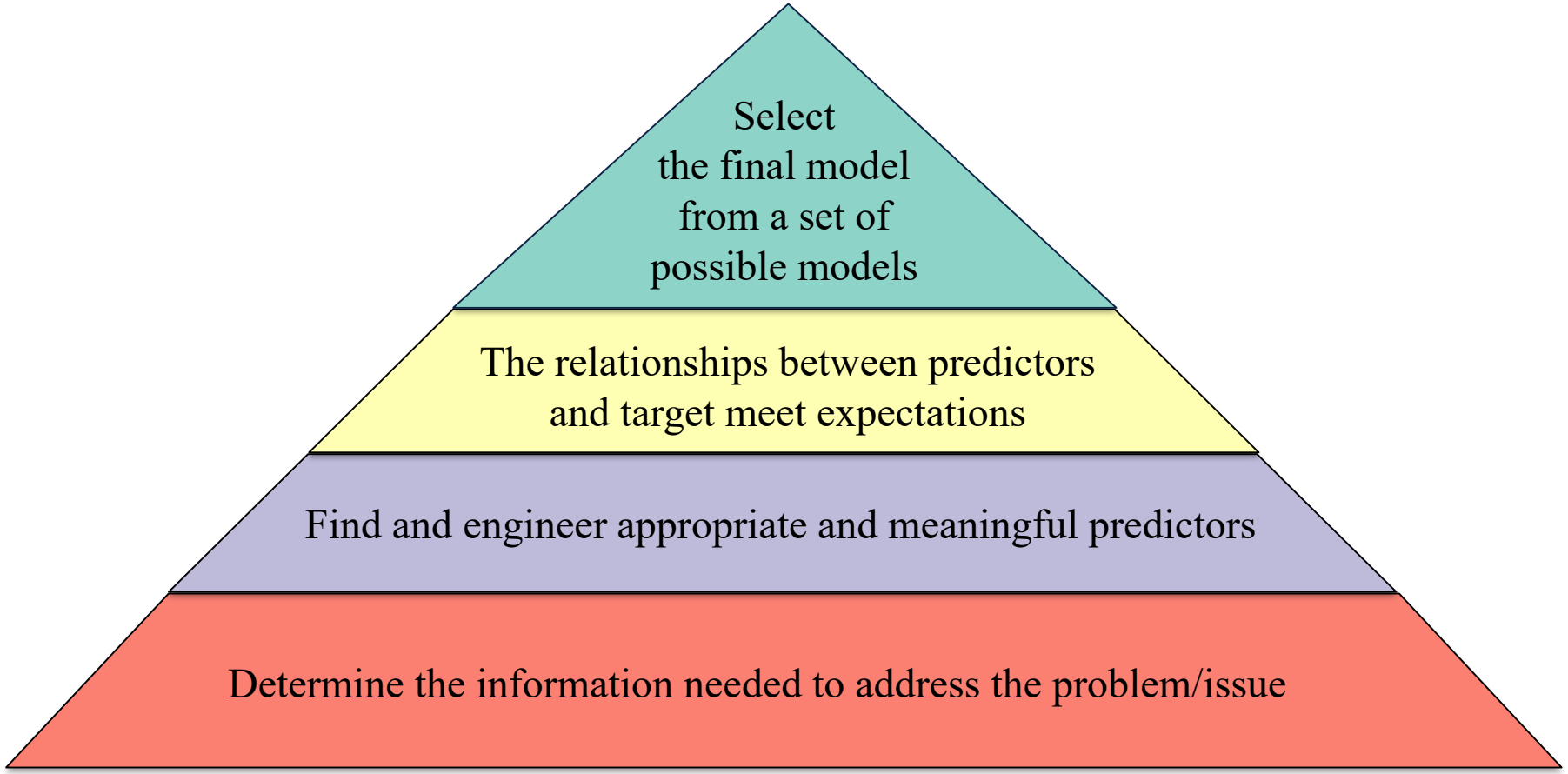
A REAL WORLD APPLICATION OF PREDICTIVE ANALYTICS

alteryx | The Thrill of Solving



- The predictive analytics pyramid
- A hands application to predict electricity demand for an electric utility tomorrow

THE PREDICTIVE ANALYTICS PYRAMID



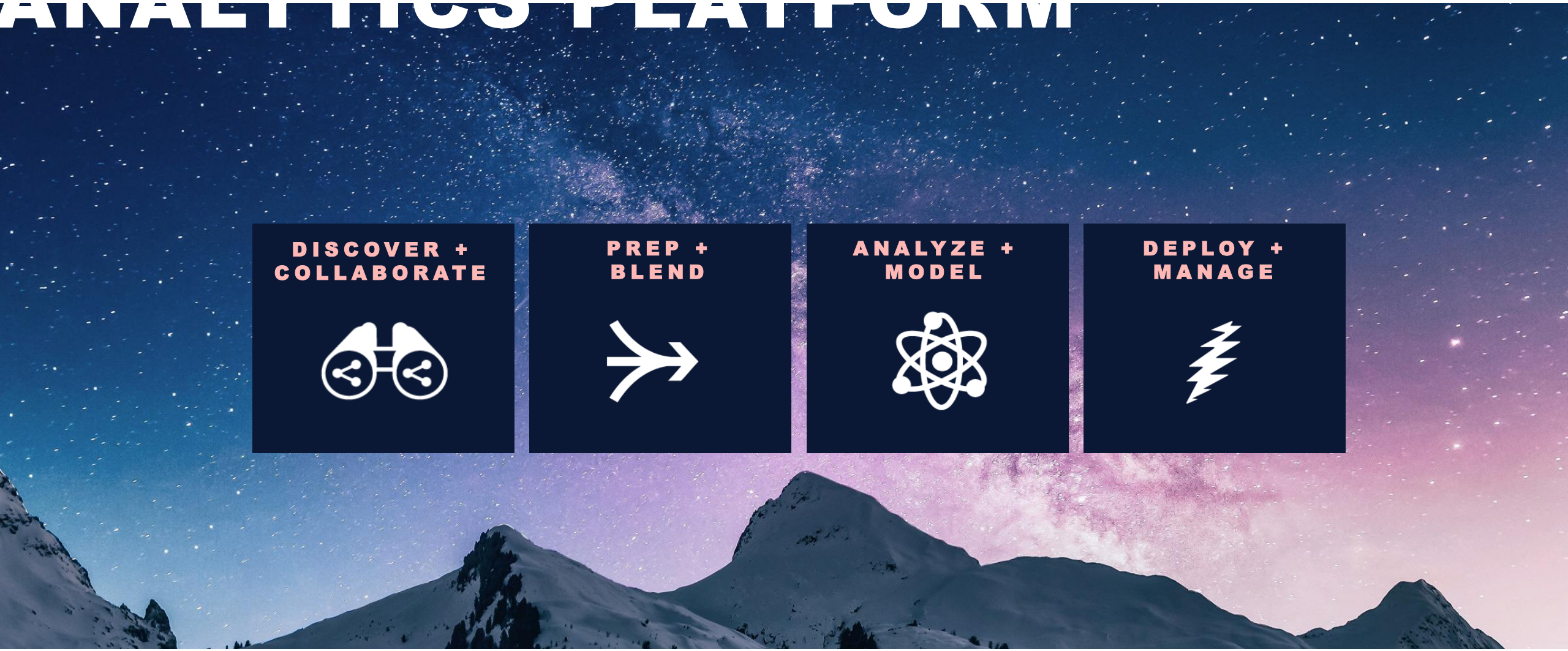
DETERMINE THE AMOUNT OF ELECTRICITY TO SUPPLY TOMORROW

- **The decision**: An electric utility needs to determine the amount of electricity it needs to produce and (potentially) buy tomorrow, and whether to sell power to other utilities
- **The information needed**: The critical information needed is the amount of electricity their customers will demand for each hour tomorrow
- **Meaningful predictor metrics**: Metrics that have proven valuable for predicting hourly electricity demand include:
  - The temperature this hour and in the preceding hour
  - The time of day, day of week, and month of year
  - The length of time since sun rise and the prior sunset
- To Alteryx, in order to provide the needed predicted information

PUTTING THE POWER OF PREDICTIVE INTO PRODUCTION

alteryx | The Thrill of Solving

FLEXIBLE END-TO-END



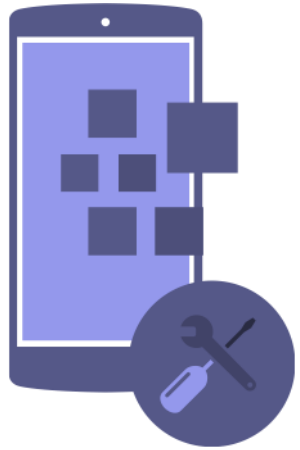
alteryx | The Thrill of Solving

MODEL DEPLOYMENT METHODS

REPORTS

INTERACTIVE DASHBOARDS

REAL-TIME APPLICATIONS

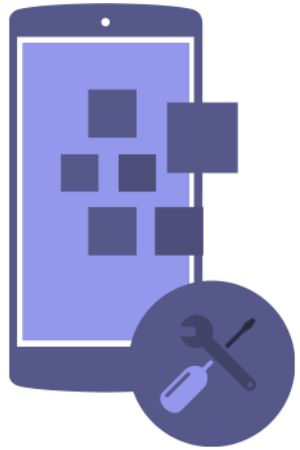


MODEL DEPLOYMENT METHODS

REPORTS

INTERACTIVE DASHBOARDS

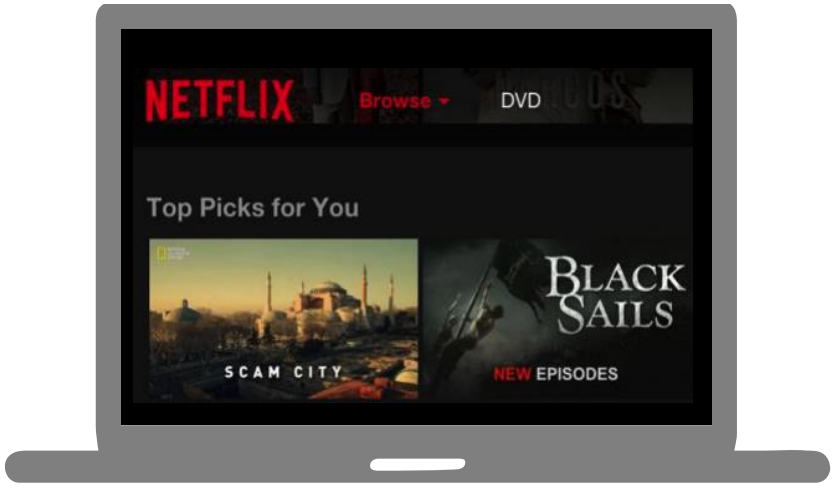
REAL-TIME APPLICATIONS



DATA SCIENCE IS ABOUT PRACTICAL, REAL-WORLD SOLUTIONS

Carl wants to watch a good movie

Hey, Carl.
Check these out!



alteryx | The Thrill of Solving

EXPLANATION ISN'T ALWAYS IMPORTANT

Sarah builds the algorithm.

Carl would like Frozen because Cindy liked it.

	Movie 1	Movie 2	Movie 3	Movie 4	Movie 5	Movie 6	Movie 7	Movie 8	Movie 9	Movie 10	...	Movie 17770
User 1			1		2							3
User 2		2		3	3			4		?		
User 3							5	3				
User 4	2				3			2				2
User 5		2		3		5		4		2		4
User 6			2									
User 7			2					4	2			
User 8	3	1			3	4		5		4		
User 9									3			
User 10			1		2							2
...												
User 480189		4			3			3				

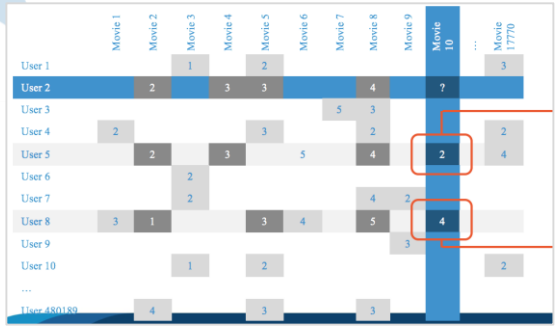
Carl

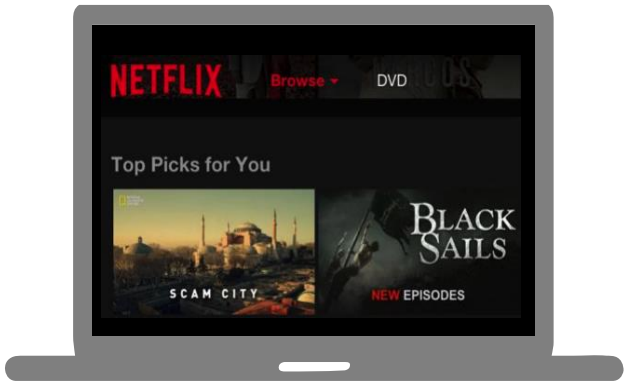
Cindy

http://courses.washington.edu/css490/2012.Winter/lecture_slides/08b_collaborative_filtering_1_r1.pdf

alteryx | The Thrill of Solving

Sarah's algorithm is seamlessly embedded





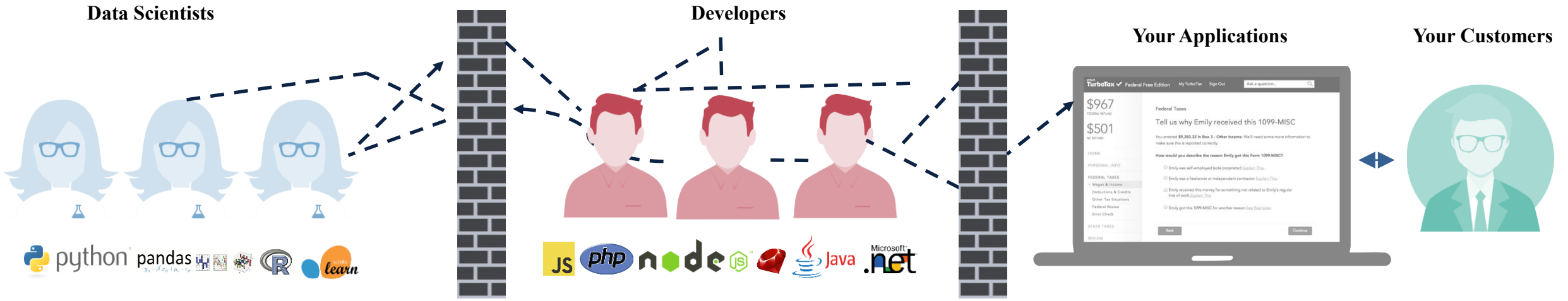
alteryx | The Thrill of Solving

MODEL DEPLOYMENT REALITY

DEPLOYMENTS TAKE 12-20 WEEKS

COST TO DEPLOY 1 MODEL RUNS IN EXCESS OF $250,000

< 10% OF MODELS MAKE IT TO PRODUCTION



Build a model in R or Python

Painstakingly rewrite models into other languages

Write *more* custom code to integrate

Customer benefits

alteryx | The Thrill of Solving

SELF-SERVICE MEETS REAL-TIME

MANAGE & DEPLOY

- Deploy R & Python models into production without recoding
- Embed machine learning models into business processes
- Manage and monitor models, version control and iterate

DYNAMIC PRICING

Suggest prices through a website or app

RECOMMENDER SYSTEMS

Use predictive models to forecast consumer behavior

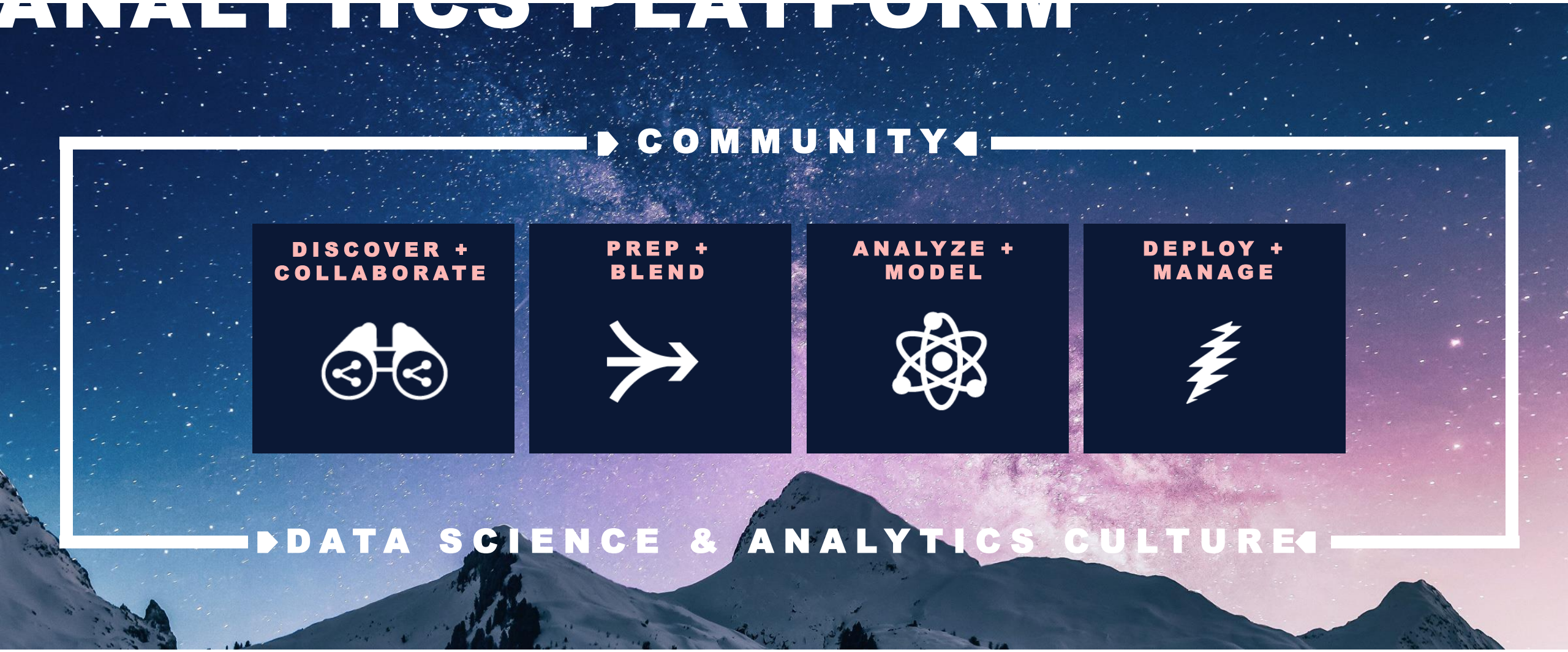
REAL-TIME CREDIT SCORING

Determine the credit approval via embedded app

FRAUD DETECTION

Automatically determine if transactions are fraudulent

FLEXIBLE END-TO-END



alteryx | The Thrill of Solving

CONCLUSION

CONSUMER PRODUCTS

Best Buy · Dick's Sporting Goods · The Home Depot · Kroger · Levi's · Nordstrom

ENERGY & UTILITIES

Ambit Energy · bp · EDF · Marathon · Shell · Toronto Hydro · Koch · World Fuel Services

FINANCIAL SERVICES

AAA · Charles Schwab · CNA · First Tennessee · HSBC · Paychex · Voya Financial · Wells Fargo

HEALTHCARE

Abbott · Anthem BlueCross BlueShield · Cardinal Health · Highmark Health Services · Humana · Roche · Quest Diagnostics

MANUFACTURING

Boeing · Brunswick · flex · Ford · GE · Ingersoll Rand · Schneider Electric · Toyota

MEDIA

Clear Channel Outdoor · Cineplex · nielsen · IPG Mediabrands · Omnicom · Sony Pictures Television · tribune direct (Direct Response. Direct Results.)

PROFESSIONAL SERVICES

accenture · acxiom · Bain & Company · BCG (The Boston Consulting Group) · Harte Hanks · Kurt Salmon · Marcus & Millichap

RESTAURANTS

Bloomin' Brands · Buffalo Wild Wings · DQ · dunkin' brands · McDonald's · Panera Bread

RETAIL

adidas · Coca-Cola · Honeywell · Johnson & Johnson · P&G · Unilever · Sally Beauty

TECHNOLOGY

hp · EMC² · Marketo · Intuit · Philips · Microsoft · VIZIO

TELECOM & CABLE

CenturyLink · Ericsson · TELUS · ViaSat · Spectrum Reach · T-Mobile · vodafone · verizon

TRAVEL & HOSPITALITY

Alaska · Carlson Rezidor Hotel Group · The Cosmopolitan of Las Vegas · Hyatt · Southwest · Vail Resorts · Virgin Atlantic · Voyages

EMBARKING ON THE PATH TO PREDICTIVE

- Predictive Analytics Templates
- Alteryx Academy Free Training in our Alteryx Community
- Alteryx Predictive Analytics Udacity Course

Thank you

Q+A