

# Pivotal

# 大数据时代:寻找数据科学家

Jimmy Chiu jchiu@pivotal.io

### The new digital customer

Rising and continuously changing expectations around experiences



Rising and continuously changing expectations around experiences

### The new digital customer



### **INTELLIGENT APPLICATIONS ARE THE NEW FACE OF BUSINESS**

### Challenges to Realizing Business Value from Your Data



### **Smart Data App**

### Increases Your Speed Through the Virtuous Cycle



## **Pivotal: Technology Leadership**



### **Data Science Goes Further Than Bl**



### **Typical Benefits of Data Science vs Traditional Analytics**

Greater accuracy through modeling detailed patterns of action rather than summarized data

- Example: ATM Cash Sufficiency
- Use large historical data to enable future prediction and hypothetical scenarios over explaining the past
- Example: Insurance policy lapse

Discovering more data points that drive an important outcome

Example: Manufacturing, Fintech

Solving a new business problem that is unsolvable with traditional analytics

Example: Customs clearance

Faster time to insights by using a big data platform, faster time to operationalization

## **Key Platform Recommendations**

- Select a platform that handles the next generation of big data providing multiple ways to query and analyze data at rest.
- Rapid software development requires an agile data infrastructure. Analytics are experimental in nature. Analytics are software. Select a platform that allows for both rapid insights as well as rapid application development. Buy one platform, not multiple solutions.
- Use In-Database Analytics and embrace open source R. Don't lock yourself into a proprietary stack!
- Build a data practice internally aligned with product innovation and software development. Engage an enablement partner if needed, not consultants.

## **Analytics Technology Stack**





### Macro-programming

- Use SQL constructs to orchestrate • parallel bulk data processing
- Much richer than the MapReduce

### **Micro-programming**

- Use User-Defined Functions in PL languages to implement local computations.
- Allows reuse of many libraries built for different languages in Greenplum.

## **Work with Pivotal Labs Agile Application Team**









#### The Vantive Business Dashboard

#### CHALLENGE

- 20 internal stockholders with varying opinions on how the app should function
- Ensuring compliance with privacy

regulations while working with a external data vendor

#### **SOLUTION**

- Discovery & Framing to focus the objective and clarify the vision of stakeholders
- Educated Visa developers and product managers on agile methods
- Designed the iPad-optimized

Dashboard screens

#### **OUTCOME**

- Delivered the iPad app in shorter timeline than expected
- The Vantive Dashboard helps millions of merchants by comparing performance with competitors and realize growth opportunities from customer purchase patterns









Pivotal **IT**太咖谚

Web Version Data Workflow Tool

#### CHALLENGE

- 1st to market with fully Android app •
- Compress timelines, agile delivery •

#### **SOLUTION**

- 4-month delivery and 50% reduction in testing cyclesAgile methodology integrated at

discrete points

#### **OUTCOME**

Record delivery time with top ratings (4-5 stars in Google Play)

Currently working on Phase 2 of the • project (re-platforming middleware stack to optimize omnichannel delivery across phones and tablets)





### **Creating competitive advantage**



Increased regulatory compliance

Reduced fraud and security risks

Improved Customer Satisfaction & Time-to-Market Optimized Business Processes & Operations

### Why build what you can buy?



### **INNOVATE RATHER THAN INTEGRATE**

## 福特Connected Car解決方案架構



- 基於車輛感測器資料、車載電腦資料和歷史資料、
   協力廠商資料,可實現如下大資料分析場景:
- 形式路徑、目標,以及可達行駛範圍
- 路況及時預警和行駛路徑建議
- 車輛異常狀況快速捕獲
- 車輛周邊協力廠商資訊服務,包括加油站、便利店、 休息站等
- 基於地理資訊的服務
- 零件維修&庫存管理

- ....



# **Getting Started**

### **Data Science Workshop**

acilitation	Data Review				
	What data exist?	Analytics Prod Use Case and problem documentation Success measures Basis of comparison	blem Definition IT Platform		
	How much exist? Is it available		How to set up a Pivotal instance Who will take ownership?	SOW Creation	

### **Data Science Labs: Packages**

Non Model Bu	uilding Labs	Model Building Labs	
LAB PRIMER (2 Week Workshop)	LAB 100 (Analytics Bundle)	LAB 600 (6-Week Lab)	LAB 1200 (12-Week Lab)
<ul> <li>Delivers a customized analytics roadmap with prioritized opportunities and architectural recommendations</li> <li>Building on Pivotal learnings and experience</li> <li>Two week light touch customer interviews plus 1- day facilitated session</li> </ul>	<ul> <li>Customer data rapid insights in 2 weeks</li> <li>Requires customer data readiness</li> <li>No model creation</li> </ul>	<ul> <li>Simple Data Science model building</li> <li>Requires customer data readiness</li> <li>Ready-to-deploy model</li> <li>Includes data load, data review, model build and handover</li> </ul>	<ul> <li>Complex Data Science model building</li> <li>Requires customer data readiness</li> <li>Ready-to-deploy model</li> <li>Includes data load, data review, model build and handover</li> </ul>

## **Example Approach: Data Science Lab 1200**

