DEVELOPING AN *ACTIONABLE* BIG DATA BUSINESS STRATEGY

BILL SCHMARZO, CTO, EMC GLOBAL SERVICES

Follow me on twitter: @SCHMARZO

© Copyright 2016 EMC Corporation. All rights reserved.
Measures degree to which organizations have integrated data and analytics into their business models.
Step 1: Understand Business Initiatives

Step 2: Identify Business Use Cases

Step 3: Define Data Requirements

Step 4: Develop Data Science Plan

Step 5: Define Technology Plan
STEP 1: UNDERSTAND BUSINESS INITIATIVES

WHAT ARE THE CHARACTERISTICS OF A BUSINESS INITIATIVE THAT MAKE IT A GOOD BIG DATA CANDIDATE?

- Sense of Urgency
- Important To The Business
- Compelling ROI
- Clear Business Ownership
- Strong IT Leadership
- Bounty of Potential Data Sources
- Analytics Friendly
EXAMPLE: [CHIPOTLE] BUSINESS INITIATIVES

Chipotle Business Initiatives

• Build people culture that attracts and empowers top performers
• Grow revenues (up 20.3% in 2012) opening new stores (opened 183 in 2012)
• Increase comparable restaurant sales growth (7.1% in 2012)
• Marketing building Chipotle brand and engaging with our customers

Dear Shareholders,

We are pleased with Chipotle’s performance in 2012, and are confident that the continuing strength of our business is a direct result of our focus on the key elements that drive our business: passionately source local and premium ingredients; together, these principles are at the heart of our mission to change the way people think about and eat fast food.

Our food culture sets us apart from other restaurants. We have always used great quality ingredients and prepared the food we serve with classic culinary techniques in open kitchens. We are proud of the way we source the finest ingredients we can find and proudly display how we work and treat our customers with great care, because we know it results in exceptional tasting food that our customers appreciate. We believe this, as well as our sit-down style of service, are the reasons more people talk about where they ate than where they ordered from, and they discover the special way we serve our food. We believe the more people care about their food, and share it with others, the more likely they are to become long-term customers of Chipotle.

Throughout 2012, we continued to push ourselves to find better, more sustainable sources for our ingredients—we used and refined our cooking techniques in order to minimize oil use, and we eliminated the use of onions, and now servemade with rice that is grown on our family’s farm in the United States. Finally, we reinforced our commitment to the responsible treatment of animals that are raised in our restaurants. We were featured in the cross-town, a respected food publication, for our commitment to sourcing and serving high-quality ingredients.

We continue to build a culture that attracts and empowers top performers. Our core management team, which is made up of former top performers, is one of the most respected in the industry. In 2012, we opened 140 new restaurants, giving us a total of 975 in-store employees, making it the largest Hispanic-owned company in the United States. We are seeing a higher percentage of candidates interviewed for top management positions that were previously filled by our Hispanic staff. We are very proud of the fact that our Hispanic employees are the majority of our management team, and that the quality of managers in our restaurants is getting better all the time. At our September All Manager Conference, we brought together an incredible group of 8,000 leaders to share insights about our vision, and to discover opportunities for them to learn about best practices in their stores and help them run their restaurants even better. This ongoing experience is what we continue to consider as one of the reasons for the strength of our restaurants.

Last year we opened 140 restaurants, grew our revenue by 25.3% to $5.73 billion, and saw comparable restaurant sales grow 7.1% for the year. Our restaurant-level metrics were among the highest in the industry at 2.1%. We are proud of our performance, and anticipate significant growth and success. We plan to open an additional 150-160 restaurants in 2013, and we anticipate that we will continue to grow at a rate that we can fund organically and not have to rely on debt or equity capital. We have a strong balance sheet and cash flow, and are confident that we can achieve our goals in the future.

动力来自于我们对Chipotle品牌的承诺和与客户建立联系。我们认识到，这是一种强大的、深刻的关系，可能是唯一可能的"可持续"的，为我们的企业，我们的品牌，我们的产品，以及我们的客户，创造价值。我们将继续努力，与我们的客户建立牢固的关系，同时保持我们的高标准。随着我们的发展，我们将继续关注我们的员工，与他们建立一种信任和尊重的关系，这将有助于我们实现我们的目标。
STEP 2: IDENTIFY BUSINESS USE CASES
WHERE AND HOW DO YOU INTEGRATE DATA AND ANALYTICS INTO THE ORGANIZATION’S BUSINESS STRATEGY?

• Identify business **Use Cases** that support to key **Business Initiatives**

• Identify **Business Stakeholders** who either impact or are impacted by the targeted business initiative

• Leverage **Prioritization Matrix** to drive organizational consensus around use cases
EXAMPLE: [CHIPOTLE] USE CASE

INCREASE STORE TRAFFIC: Leverage local events (little league games, soccer tournaments, high school sports, college sports, concerts, etc.) with more localized marketing campaigns, promotions and sponsorship to increase store traffic during these local events.

BUSINESS POTENTIAL

• Develop highly-localized marketing campaigns and promotions based on the local store demographics, economics and social vibrancy
• Understand local events’ demographics in order to make sponsorship and promotional determinations

IMPLEMENTATION RISKS

• Tracking effectiveness of campaigns
• Capturing (screen scraping) local events data
• Ability to create localized campaigns

IMPACT CHIPOTLE FINANCIAL GOALS

<table>
<thead>
<tr>
<th></th>
<th>Store Revenue Growth</th>
<th>Customer Acquisition</th>
<th>Customer Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market Basket Margin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product Cross-Sell</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Products</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Overall Impact on Goals: 2.0

IMPLEMENTATION CONSIDERATIONS

• Capturing and governing additional data
• Ability to monitor sponsorship effectiveness
• Store manager training and capabilities

Overall Ease of Implementation: 2.5

Revenue Potential: $_______
**EXAMPLE: [CHIPOTLE] BUSINESS STAKEHOLDER**

**BUSINESS INITIATIVE:** INCREASE SAME STORE SALES

**Store Manager:** Typically 2 to 4 years of Chipotle experience; promoted through the ranks to store manager within same store

**Quote:** “I love Chipotle, so working here only seemed natural”

<table>
<thead>
<tr>
<th>Key Decisions</th>
<th>Supporting Business Questions</th>
<th>Pain Points</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Staffing</strong></td>
<td>• What’s the likely store traffic throughout the day?</td>
<td>Lacks real-time model that allows him to update changing demand variables</td>
</tr>
<tr>
<td></td>
<td>• When and for how long are the local events?</td>
<td></td>
</tr>
<tr>
<td><strong>Inventory</strong></td>
<td>• Given store traffic estimates, how much additional inventory will I at what points in the day?</td>
<td>Doesn’t have insights into extra inventory at neighboring stores that might fill temporary inventory gaps</td>
</tr>
<tr>
<td></td>
<td>• What local events that could impact inventory?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• What are the types of local events?</td>
<td></td>
</tr>
<tr>
<td><strong>Customer Satisfaction</strong></td>
<td>• What’s the best comp to give unsatisfied customer?</td>
<td>Doesn’t know “how valuable” that particular customer is to the Chipotle</td>
</tr>
<tr>
<td></td>
<td>• How important is that customer to me?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• What is the cost of the different comp options?</td>
<td></td>
</tr>
<tr>
<td><strong>Production</strong></td>
<td>• Do I need to shift staff to produce more materials given the store traffic and time of day?</td>
<td>Doesn’t have insights into local events that could dramatically impact production needs</td>
</tr>
<tr>
<td></td>
<td>• Which products are most popular for local events?</td>
<td></td>
</tr>
</tbody>
</table>
EXAMPLE: [CHIPOTLE] KEY DECISIONS

BUSINESS INITIATIVE: INCREASE SAME STORE SALES

STAKEHOLDER: STORE MANAGER

STAFFING
How many staff do I need? What skills do I need? What roles do I need?

SCHEDULING
When (dates and times)? For how long? How much overtime?

INVENTORY

PRODUCTION
How much to produce? When? What types of items to produce?

CORPORATE SALES (CATERING)
Who to target? When are their events? What’s the priority targets?

LOCAL EVENTS PROMOTIONS
Which events? What type of promotion? When to promote?

LOCAL EVENTS SPONSORSHIPS
Which events? What type of sponsorship? When to sponsor?
EXAMPLE: [CHIPOTLE] PRIORITIZATION

INCREASE SAME STORE SALES

Use Cases

- **A** Increase Store Traffic
- **B** Increase Shopping Basket Size
- **C** Increase Corporate Events
- **D** Improve New Product Introductions
- **E** Improve Promotional Effectiveness
- **F** Increase Special Events
STEP 3: DEFINE DATA REQUIREMENTS
WHAT DATA DO YOU NEED TO SUPPORT THE BUSINESS USE CASES?

• **Brainstorm** the variables and metrics (data sources) that *might* yield better predictors of business performance

• **Assess** each data source from the perspectives of **Business Value** and **Implementation Feasibility**
EXAMPLE: [CHIPOTLE] DATA SOURCES

✓ Point of Sales Transactions
✓ Market Baskets
✓ Product Master
✓ Store Demographics
✓ Competitive Stores Sales
✓ Store Manager Notes
✓ Employee Demographics
✓ Store Manager Demographics
✓ Consumer Comments
✓ Weather
✓ Traffic Patterns
✓ Yelp
✓ Zillow/Realtor.com
✓ Twitter/Facebook/Instagram
✓ Twellow/Twellowhood
✓ Zip Code Demographics
✓ EventBrite
✓ MaxPreps
✓ Mobile App
✓ ...

© Copyright 2016 EMC Corporation. All rights reserved.
## EXAMPLE: [CHIPOTLE] DATA ASSESSMENT

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Increase Store Traffic</th>
<th>Increase Shopping Bag Revenue</th>
<th>Increase # Corporate Events</th>
<th>Increase Promotional Effectiveness</th>
<th>Improve NPI Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point of Sales Transactions</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Market Baskets</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Store Demographics (Zip Code)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Local Competitive Stores</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Store Manager Demographics</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Consumer Comments</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Social Media</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Weather</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Local Events</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Traffic</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Zillow</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

### BUSINESS VALUE

### IMPLEMENTATION FEASIBILITY

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Ease of Acquiring</th>
<th>Cleanliness</th>
<th>Accuracy</th>
<th>Granularity</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Point of Sales Transactions</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Market Baskets</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Store Demographics (Zip Codes)</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Competitive Stores Sales</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Store Manager Demographics</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Consumer Comments</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Social Media</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Weather</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Local Events</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Traffic</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>Zillow</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>
STEP 4: DEVELOP DATA SCIENCE PLAN

WHAT ARE THE DATA SCIENCE REQUIREMENTS NECESSARY TO SUPPORT THE BUSINESS PLAN?

 ✓ People (Tiger Team)
 ✓ Capabilities (Tools)
 ✓ Process
DATA SCIENCE TIGER TEAM PLAYERS
THE DATA SCIENCE “TIGER TEAM” NEEDS THE FOLLOWING ROLES:

- Business Subject Matter Expert (SME)
- Data Czar
- Data Engineer
- Data Architect
- Data Scientists
- Data Visualization
- Data Governance
- App Dev/User Experience
DATA SCIENCE CAPABILITIES

THE DATA SCIENCE PLAN NEEDS TO ADDRESS THE FOLLOWING DATA SCIENCE CAPABILITIES (TOOLS):

- Data acquisition and ingest
- Data repository
- Data discovery
- Data transformation and enrichment
- Data visualization (Tableau, DataRPM, ggplot2)
- Analytic modeling (SAS, R, MADlib, Statistica, H20)
DATA SCIENCE PROCESS

Step 1: Define Hypothesis to test or Prediction to be made

Step 2: Gather data...and more data (Data Lake: SQL + Hadoop)

Step 3: Prepare the data (transformation, enrichment)

Step 4: Visualize the data (Tableau, DataRPM, ggplot2, ...)

Step 5: Build analytic models (SAS, R, MADlib, Mahout, ...)

Step 6: Evaluate model results (probabilities, confidence levels)
STEP 5: DEVELOP TECHNOLOGY PLAN

WITH THE BUSINESS USE CASES, DATA AND DATA SCIENCE REQUIREMENTS IN HAND, WE ARE NOW READY FOR:

DATA ARCHITECTURE  BIG DATA ECOSYSTEM
DATA ARCHITECTURE

HADOOP DISTRIBUTION

DATA CURATOR
- INGEST
- INDEX
- ENRICH

PLATFORM MANAGER
- ADMINISTRATION
- ANALYTICS CATALOG
- DATA CATALOG

PRIVATE CLOUD
- COMPUTE
- HYPERSCALE
- DATA LAKE
- REAL-TIME STREAMING

DATA SCIENCE
- DATA EXPLORATION
- DATA ENRICHMENT
- DATA VISUALIZATION
- ANALYTIC MODELING

DATA GOVERNOR
- LINEAGE
- QUALITY
- SECURITY
WHAT SHOULD YOU DO NOW?
WHAT ARE THE SETS OF ACTIONS AND IDEAS THAT YOU CAN TAKE BACK TO YOUR OFFICES BASED UPON TODAY’S SESSION?

✓ Identify targeted **business initiative**
✓ Identify supporting **business use cases**
✓ Identify key **business stakeholders**
✓ Capture key **business decisions**
✓ Identify and brainstorm **data** to explore
✓ Identify **architecture** and **technology** requirements
THANK YOU!


To Achieve Big Data’s Potential, Get It Into The Boardroom
http://www.entrepreneur.com/article/236125

Big Data Business Model Maturity Index (animation)

Big Data Business Model Maturity Index Guide

Keys to Big Data Success? KISS!

How I’ve Learned To Stop Worrying And Love The Data Lake

Contact Information
Email: William.Schmarzo@emc.com
Twitter: @schmarzo