All Meat, No Fillers

Rob Ortiz
Division Director
Keyence

Brian Neely
Sr. Data Scientist
Keyence

#data19
Rob Ortiz

Started online retailer at age 19

16 years with Keyence

Launched deployed analytics team

Current director of Keyence AI

Aspiring Tableau Jedi

Writes sci-fi novels

Builds board games

#data19
Founded in 1974

Automation company

Forbes 100 Most Innovative Companies

“Most important company, you’ve never heard of”
—Financial Times

5–7 Years

#data19
The job of the data scientist is to ask the right questions.

Hilary Mason
GM of Machine Learning, Cloudera
“Hey Rob, quick question. I got alerted by the Tableau this morning that sales volume dropped below the threshold we set. Do you have time for a quick meeting to help figure out why?”

“Sure, let me just go grab my laptop.”
“As I mentioned, sales volume is down. Any ideas?”

“Well my assumption is something to do with the new product we launched. Do you have a dashboard for that yet?”

“Alright, well do you have some ideas where to start?”

“Not yet, no. I can build it pretty quickly though.”

“Terry, this is why we use Tableau. It’s explorer functions let you dig into the data.”

“No, but the guy who built this presentation was limited in images to choose from.”
“Yes please, oh also let’s check deliveries. It might be that as well. And while I’m thinking about it, can we look by salesperson?

“Ok, Got it. I’ll get to work on it.”
A few hours later
Dashboards are done

Hmm, this doesn’t quite show what I need. Can you let me drill down by sub-group?

Hang on, I’ll add it

Also, had another thought. Can we see sales for new salespeople only? I think we may have an issue with tenure

Sure thing, I’ll get right on it
A few days later
So after doing some analysis we’re pretty sure the issue is with X.

Well now that we have it narrowed down to a few options, let me get the Data Science team on it.

I don’t see it being X. I still think Y is the bigger issue.

I think we need to see this from a different angle. What if we looked at Z?
Why Do We Get Stuck into this Cycle?

As Data Analysts it is our job to build the deliverables

We **can** make the most all powerful “Genie”

In a Domain Expert, Zen Master, AI….

We **can** build the most beautiful “Lamp”

But at the end of the day, it is the skill of the “Wisher”

They must ask the right questions to lead to results
3-6 Weeks

15-30 Minutes
Asking the right questions

How does price impact repeat?

Which Salesperson has the best/worst repeat?

Which products aren’t repeating?

Is it the sales teams fault?

How does industry impact our repeat rate?

Is it the operations teams fault?

Which customers who usually repeat, aren’t?
Feature Engineering

The process of using **domain knowledge** of the data to create new angles and viewpoints that will make the insights in the data appear more sharply.

But what happens if you don’t have **domain knowledge**?
Asking the Right Questions

The first question:
• What is the most important information to show?

The second question:
• What can we change?

The third question:
• What has the largest impact, for the least amount of effort?
So we started building
Warning: Data Science Ahead
Making of the Magic

Automated feature engineering/learning:

• Builds meta data
• Generates aggregates
• Common calculations

Gradient boosted decision tree:

• Predicts a target label and ranks likely outcomes
• Handles nulls/missing values well
Warning: Data Science Ahead
<table>
<thead>
<tr>
<th>Date Time</th>
<th>Customer ID</th>
<th>Product Group</th>
<th>Qty</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019/11/04 (MON) 12:30</td>
<td>10111</td>
<td>Sensor</td>
<td>5</td>
<td>$149.95</td>
</tr>
<tr>
<td>2019/11/04 (MON) 14:21</td>
<td>10135</td>
<td>Measurement</td>
<td>9</td>
<td>$891.00</td>
</tr>
<tr>
<td>2019/11/06 (WED) 19:38</td>
<td>10352</td>
<td>Software</td>
<td>1</td>
<td>$1,899.99</td>
</tr>
<tr>
<td>2019/11/09 (SAT) 19:30</td>
<td>10111</td>
<td>Sensor</td>
<td>3</td>
<td>$199.96</td>
</tr>
<tr>
<td>2019/11/10 (SUN) 11:27</td>
<td>10165</td>
<td>Measurement</td>
<td>1</td>
<td>$99.99</td>
</tr>
<tr>
<td>2019/11/10 (SUN) 19.30</td>
<td>10891</td>
<td>Sensor</td>
<td>1</td>
<td>$29.49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Customer ID</th>
<th>Product Group</th>
<th>Sum Qty</th>
<th>Avg Qty</th>
<th>Max Qty</th>
<th>Min Qty</th>
<th>Sum of Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>10111</td>
<td>Sensor</td>
<td>8</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>$349.91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Day of Week</th>
<th>Count unique customers</th>
<th>Sum Sales</th>
<th>Sum QTY</th>
<th>Sale per unit</th>
<th>Sale per customer</th>
</tr>
</thead>
<tbody>
<tr>
<td>MON</td>
<td>2</td>
<td>$1,040.95</td>
<td>14</td>
<td>$74.35</td>
<td>$520.48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Product Group</th>
<th>Count transactions</th>
<th>Sale per unit</th>
<th>Sale per transaction</th>
<th>Count of Weekday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor</td>
<td>3</td>
<td>$42.16</td>
<td>$126.47</td>
<td>1</td>
</tr>
</tbody>
</table>
**Transactions**

<table>
<thead>
<tr>
<th>Date Time</th>
<th>Customer ID</th>
<th>Product Group</th>
<th>Qty</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019/11/04 (MON) 12:30</td>
<td>10111</td>
<td>Sensor</td>
<td>5</td>
<td>$149.95</td>
</tr>
<tr>
<td>2019/11/04 (MON) 14:21</td>
<td>10135</td>
<td>Measurement</td>
<td>9</td>
<td>$891.00</td>
</tr>
<tr>
<td>2019/11/06 (WED) 19:38</td>
<td>10352</td>
<td>Software</td>
<td>1</td>
<td>$1,899.99</td>
</tr>
<tr>
<td>2019/11/09 (SAT) 19:30</td>
<td>10111</td>
<td>Sensor</td>
<td>3</td>
<td>$199.96</td>
</tr>
<tr>
<td>2019/11/10 (SUN) 11:27</td>
<td>10165</td>
<td>Measurement</td>
<td>1</td>
<td>$99.99</td>
</tr>
<tr>
<td>2019/11/10 (SUN) 19.30</td>
<td>10891</td>
<td>Sensor</td>
<td>1</td>
<td>$29.49</td>
</tr>
</tbody>
</table>

**Customer**

<table>
<thead>
<tr>
<th>Customer ID</th>
<th>Industry</th>
<th>Size</th>
<th>Salesperson</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>10009</td>
<td>Retail</td>
<td>$100 M</td>
<td>Ian</td>
<td>North</td>
</tr>
<tr>
<td>10087</td>
<td>Electronics</td>
<td>$1 B</td>
<td>Cathy</td>
<td>South</td>
</tr>
<tr>
<td>10111</td>
<td>Home Goods</td>
<td>$5 B</td>
<td>Ian</td>
<td>East</td>
</tr>
<tr>
<td>10178</td>
<td>Retail</td>
<td>$250 M</td>
<td>Blake</td>
<td>East</td>
</tr>
<tr>
<td>10265</td>
<td>Wholesale</td>
<td>$10 B</td>
<td>Cathy</td>
<td>South</td>
</tr>
<tr>
<td>10352</td>
<td>Electronics</td>
<td>$100 M</td>
<td>Ed</td>
<td>South</td>
</tr>
</tbody>
</table>

**Web activity**

<table>
<thead>
<tr>
<th>Customer ID</th>
<th>Date Time</th>
<th>Access</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>10009</td>
<td>2019/11/02 (SAT) 8:32</td>
<td>Help Guide</td>
<td>Google</td>
</tr>
<tr>
<td>10009</td>
<td>2019/11/03 (SUN) 14:18</td>
<td>Account Info</td>
<td>Direct</td>
</tr>
<tr>
<td>10009</td>
<td>2019/11/05 (TUE) 20:53</td>
<td>Help Guide</td>
<td>Email</td>
</tr>
<tr>
<td>10111</td>
<td>2019/11/01 (FRI) 11:10</td>
<td>Video</td>
<td>LinkedIn</td>
</tr>
<tr>
<td>10265</td>
<td>2019/11/06 (WED) 15:37</td>
<td>Account Info</td>
<td>Email</td>
</tr>
<tr>
<td>10265</td>
<td>2019/11/01 (FRI) 14:21</td>
<td>Account Info</td>
<td>Direct</td>
</tr>
</tbody>
</table>

**Sales activity**

<table>
<thead>
<tr>
<th>Salesperson</th>
<th>Date Time</th>
<th>Event</th>
<th>Customer ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blake</td>
<td>2019/11/01 (FRI) 11:18</td>
<td>Phonecall</td>
<td>30823</td>
</tr>
<tr>
<td>Cathy</td>
<td>2019/11/04 (MON) 16:14</td>
<td>Salescall</td>
<td>10087</td>
</tr>
<tr>
<td>Donna</td>
<td>2019/11/04 (MON) 16:18</td>
<td>Phonecall</td>
<td>15765</td>
</tr>
<tr>
<td>Ed</td>
<td>2019/11/05 (TUE) 9:54</td>
<td>Web</td>
<td>19876</td>
</tr>
<tr>
<td>Blake</td>
<td>2019/11/06 (WED) 10:19</td>
<td>Service</td>
<td>21548</td>
</tr>
<tr>
<td>Cathy</td>
<td>2019/11/05 (TUE) 14:51</td>
<td>Web</td>
<td>13767</td>
</tr>
</tbody>
</table>
Feature Engineering
Retail customers
+ Purchased from Ian

Customers who purchased more than 6 times
+ But never purchased Sensors
+ And never spends more than $39.99

Company size is less than $100 M
+ Have not had a phonecall in 41+ Days

5.3%
4.9%
14.7%
So we started building

Step 1:
• Scan data for relevant information to the question

Step 2:
• Create new data from the original data
• Automatic aggregations
• Automatic calculations

Step 3:
• Combine data together in nearly infinite configurations

Step 4:
• Scan for largest Impact factors and simulate result
Asking the right questions

- How does price impact repeat?
- Which Salesperson has the best/worst repeat?
- Which products aren’t repeating?
- How are customers in the $100M range without phonecalls in the last 41 days doing?
- Is it the sales team’s fault?
- How does industry impact our repeat rate?
- Is it the operations team’s fault?
- Which customers who usually repeat, aren’t?
Increase phonecalls to $100 M companies, alert if greater than 41 days

Impact: Increase XM repeat by 14.7%
Some key points

In 2008 Keyence began building Ki
  • By 2011 it was deployed into the entire organization

Designed for the business level user, to ask “Why”?
  • And allow them to come prepared with shortcuts to support BI and DS teams

Since 2011 Keyence has seen 318% growth
  • All with only an 8.5% head count increase
Human resources story time!
Hey Rob, quick question!

Never mind, I figured it out

Hey Rob! The SQL servers down again!

I noticed NPS has dropped

Oh, that’s why! Let me go talk to them

I need a strategy to increase sales

Well that was fast!
Adding AI to your BI, C I Can Do it Too!

Thursday | 2:15pm – 3:15pm | South Pacific B
Please complete the session survey in the mobile app

View ‘My Evaluations’ in the menu or find your session under ‘Schedule’
Thank You

@KeyenceInsights

Booth 1125

linkedin.com/in/rob-ortiz-57533b22