

Charity Shows That You Don't Need a Big Budget to Succeed With Predictive Analytics

Published: 17 September 2012

Analyst(s): Joao Tapadinhas

A U.K. charity has shown how the smallest organization can achieve striking results with business intelligence, winning Gartner's BI Excellence Award 2012. We give some best practices for IT leaders wanting to use data mining and predictive analytics on a small budget.

Key Challenges

- Medway Youth Trust (MYT) had a vast amount of data about young people. It decided to experiment with business intelligence (BI) and text analytics to see whether more could be done with the data and whether this could enhance its interventional work with young people.¹
- The figures for young people not in education, employment or training (NEET) in MYT's coverage area were rising. MYT wanted to discover why. It decided to use predictive analytics to try to identify the young people most at risk of becoming NEET.
- MYT faced an 18% reduction in income from government contracts and, consequently, a 16% reduction in its full-time-equivalent (FTE) head count. It saw predictive analytics as a way to concentrate the efforts of its staff and its limited resources on those young people most likely to need help to prevent them becoming NEET.

Recommendations

- Foster a culture of innovation.
- Set clear aims to enable success to be measured.
- Have a good knowledge and understanding of the data being collected and the data to be used in the model, and put in place a governance framework.
- Reassure staff that the system won't replace their professional judgment.
- Persevere and allow the project to evolve.
- Feed analytics results into the organization's core systems, linking BI with business processes.

Analysis

Organizations hold an ever-growing amount of data. But, often, they fail to unlock the value of this data. This is particularly so among smaller organizations with limited funds and little or no experience of business analytics. Although the use of predictive analytics by commercial organizations is common to predict customer habits, such as who will buy what, MYT is one of the first organizations to use it to help people, rather than to generate revenue and profit (see Note 1). It has also shown that smaller organizations can obtain striking results from such a system and that this activity doesn't require a large budget — its system identified those most at risk of becoming NEET, enabling MYT to meet them 280 days earlier than would have been possible otherwise. It achieved this with a budget of only £123,000.

Best Practices

This report describes the best practices that other organizations, particularly small and midsize organizations, can draw from the experience of MYT with data mining and predictive analytics (see the MYT's Project Approach section). These best practices include:

- Foster a culture of innovation.
- Set clear aims to enable success to be measured.
- Have a good knowledge and understanding of the data being collected and the data to be used in the model, and put in place a governance framework.
- Reassure staff that the system won't replace their professional judgment.
- Persevere and allow the project to evolve.
- Feed analytics results into the organization's core systems, linking BI with business processes.

Foster a Culture of Innovation

Traditional ways of working can become entrenched in organizations, stifling creativity and blocking the consideration of new approaches. But MYT as a new charity, established only in April 2008, has shown what can be achieved by an organization that doesn't feel bound by history and that has created a culture of innovation. It is constantly looking at ways to solve deep-rooted social problems, rather than to perpetuate the way it delivers its services. It is focused on the end result it's trying to achieve, what it needs to do to get there and how it might approach things differently to reach that endpoint. Its strategy document, for example, is called "Starting at the End."

MYT's desired result was to prevent young people in the Medway area from becoming NEET. The organization wanted to have a much deeper understanding of what changes in young people's lives to cause them to become disengaged from society. It wanted to know why one young person is in education, employment or training, but the young person who lives next door is not.

Gary Seaman, MYT's data quality manager, saw that the charity had a mass of data and wondered whether MYT understood the data fully and whether the data could tell the organization things in a different way. In particular, Mr. Seaman wanted a piece of software that could read the text about

young people entered by MYT's professionally qualified advisers in free text fields in its database and identify any anomalies in the text against flags that record the young people's status in terms of being in work, education or training. As MYT, under its chief executive Graham Clewes, has a culture of innovation, Mr. Seaman suggested using data mining, including text analytics techniques, and MYT gave the project the go ahead.

MYT obtained structured data from the work that gave it information such as the schools the young people had attended and where they lived, by gender and ethnicity, as well as unstructured text data from its staff's interventions with young people. Using text analytics capabilities, it developed a prediction model that enabled it to identify factors — often gleaned from the unstructured data content — that contributed to young people becoming unemployed. It used the rule set generated by its prediction model to forecast which of the young people in their last year of statutory education were likely to become NEET.

MYT embarked on the project despite being a small organization with a limited budget and no experience of predictive analytics or BI. The project team consisted of its sponsor (Mr. Clewes), a member of IT (Mr. Seaman), MYT's business manager and an IBM consultant who was on-site for 27 days to help develop the model. MYT's experience shows what can be achieved when an organization listens to the ideas of staff members.

Set Clear Aims to Enable Success to be Measured

If a data mining and BI project doesn't have clear goals, the organization won't know how successful it has been. MYT's aims included:

- Unlocking the value of its data.
- Reducing the number of young people leaving education to become NEET.
- Intervening in those young people's lives at an earlier stage.

MYT had set clear aims, so it was able to measure its success against those targets. It:

- Analyzed its data, specifically the 175 million text words recorded in young people's client files, to predict which young people were most likely to become NEET.
- Targeted the 732 young people identified by its prediction model, out of a total of 3,500, as 60% or more likely to become NEET. Out of those 52% had positive outcomes and went into further education, training or employment.
- Met the young people identified as at risk of becoming NEET approximately 280 days earlier than would have been possible otherwise.

Another result of the project is MYT's success in securing a £7 million contract from local government to prevent young people from becoming NEET.

Have a Good Knowledge and Understanding of the Data Being Collected and the Data to Be Used in the Model, and put in Place a Governance Framework

An organization using predictive analytics should know and understand the data that sits in its database tables. It needs to have a really good knowledge of what data is being collected and why, and what data is to be used in the model. During the project, MYT looked at some of its data and found that either there wasn't enough data to make an impact in terms of the modeling or predictions, or that the data was so inaccurate it would have skewed the model. The business understanding side of the modeling process is key, as is the data understanding side.

A governance framework is also crucial. MYT is part of a local-area information governance framework that involves the police, the health service, education providers, almost all public sector agencies and one or two other charities. MYT operates under a clear information governance and data-sharing protocol framework that's legally robust and owned by all the relevant partners that are likely to share data. All the partners hold each other to account.

MYT has confidentiality agreements in place with all the young people with whom it engages. These agreements set out how MYT might use their data. Every time a member of staff engages with a young person, the staff member refers to the agreement, so that young people continually have a chance to change their mind about how their data is used. MYT also asked IBM to sign nondisclosure and privacy agreements.

Reassure Staff That the System Won't Replace Their Professional Judgment

It's important staff understand that a BI program enhances their work, rather than threatens it. MYT's own staff and volunteers input structured and unstructured data about young people into the organization's database on a daily basis. Its personal advisers' experience, training and professional qualifications are in working with young people. They were concerned that the system would tell them how to do their job.

MYT told its staff that the data they input would go through a modeling process to identify almost straightaway the young people most likely to need help, so that they could concentrate their efforts on them. MYT explained that the purpose of the analytics system was to inform the decisions they make, not to replace their professional judgment. It said that MYT would use the system to test their professional judgment, and it would use their professional judgment to test the system, on an ongoing basis, reviewing it every six weeks through its caseload management system.

It also explained to staff that the modeling process was an opportunity for them as professionals to test their own hypotheses as part of MYT's culture of innovation. For example, if the professional instinct of members of staff was that young people's failure to gain the right qualifications at the age of 16 was a major factor in them becoming NEET, MYT would focus on that. MYT has begun to test some of its staff's hypotheses using the modeling process, which staff have found professionally stimulating. It wasn't easy at the beginning because of professional fear, but MYT reassured its staff and showed them how analytics can add value to their work.

Persevere and Allow the Project to Evolve

Acknowledge that it isn't always going to be a smooth path — and encourage those doing the work to persevere. Be open to the prospect of the project evolving. There's always a temptation for any organization to have the perfect end goal in sight before it decides to do something. But MYT is now doing things that it didn't think about before the project began. It didn't have its current state as its goal. It started the project as an inward-looking process to improve its data quality and outcomes for young people. It first wanted to capture structured information from free text to identify the characteristics of those who have become NEET. As it did this, it saw the opportunity to take that structured information and use it in a predictive analytics model to identify those most likely to become NEET.

By allowing the project to evolve, MYT has achieved more than it set out to accomplish. It has decided to invest more money in the area to turn it into something from which the charity can derive new funds to spend on its work with young people. It has set up a subsidiary social enterprise that offers data mining, propensity modeling and data analytics services. It aims to take its innovative concept to local authorities, other charities and voluntary-sector organizations.

MYT is looking at the scope for applying the processes it has developed and the data dictionary it has built to other social problems, such as teenage pregnancy, offending behavior and families that are time-consuming for local authorities. It is also considering using a similar modeling process to identify those young people better suited to different types of apprenticeships, or better suited to different types of college courses, so that they'll be more likely to progress and finish. This is possible because MYT has a wealth of historical information about young people who haven't succeeded, so it can build an effective model to identify what will best suit the young people coming through the system.

Feed Analytics Results Into the Organization's Core Systems, Linking BI With Business Processes

MYT has embedded its risk of NEET indicator into its core system's data-entry screens, so that staff can see whether young people's scores are going down, staying static or increasing. For staff, it's rewarding to know that even if a young person is still not in work, his or her likelihood of obtaining work is increasing — they can see it on their screen, and it's scientific and objective. They can see a correlation between their intervention and making progress toward the outcome for each individual child, which they've never been able to see before. If the system shows that young people's risk of becoming NEET is static or increasing, MYT's staff can modify their interventions to reduce that risk. This enables MYT to extract full value from its analytics processes.

The project will also enable MYT to engage more effectively with its partners that have similar datasets. As MYT has evidence of success, its partners are willing to share their data, which could enrich data mining and analytics further. For example, the local government has provided MYT with some of its datasets about the performance of young people at school before the age of 11. Previously, MYT had information about young people from age 11 only, so this will enable it to see whether it could make interventions earlier.

MYT's Project Approach

MYT's project team consisted of its chief executive, data quality manager and business manager, plus one IBM consultant who worked with MYT for 27 days to help devise its predictive analytics model. The project cost £123,000.

The IBM consultant spoke with some of MYT's advisers to gain an understanding of young people's terminology to help build a data dictionary. The data dictionary was joined with much of the structured data to test the validity of the data and ensure there was enough volume to test for a model.

MYT used IBM's SPSS Modeler to join key datasets, mainly text that its advisers type into free text fields. It categorized some of the key words and themes coming out of the text. It became clear that the young people in the NEET group had certain characteristics that made them NEET. MYT ran the SPSS software on a laptop, rather than on a dedicated server.

The IBM software enabled MYT to connect to multiple data sources. For example, in its NEET prediction model, MYT was able to connect via Open Database Connectivity direct to the database tables and extract specific information. It was able to extract particular records for date of birth ranges, enabling it to work with 25,000 records rather than 58,000, which helped speed up the process.

MYT used Core+, a desktop application from CareerVision, as its core system to manage the operation. MYT's advisers received training and full access via a secure login that they could access from a desktop or via the Internet on a Web browser.

MYT used the rule set generated by its prediction model to forecast which of the young people in their last year of statutory education were likely to become NEET. Of a total of 3,500 young people, the model identified 732 as 60% or more likely to become NEET. MYT's advisers made contact with those young people in school and encouraged them to apply for college, employment, apprenticeships or training.

A year after deploying the model, MYT followed up on what the young people had done after leaving school. Of the young people identified as at risk of becoming NEET, 52% had gone into employment, education or training, but 48% still needed extra support. MYT revisited the model and applied a different technique to help target the 48% more effectively. Using predictive analytics enables MYT to learn from the data each time it runs the model, so that it can modify its interventions with young people.

The project has also changed how MYT captures data. Some advisers used to input much information into the free text field. Now they need write only one paragraph with key words, as the data dictionary picks out those key terms.

Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

"How BI Leaders Can Get Started With Text Analytics"

"Foster an Analytical Culture to Improve Your Business Intelligence and Performance Management Maturity Levels"

"Ten Reasons to Reach Beyond Basic Business Intelligence"

"Who's Who in Text Analytics"

Evidence

¹ The information in this report is derived from interviews by the author with Graham Clewes, chief executive of MYT, and Gary Seaman, the Trust's data quality manager. It also draws on information in MYT's submission to Gartner's BI Excellence Awards 2012.

Note 1 Medway Youth Trust

MYT is a small charity based in Chatham, Kent, U.K. Its goal is to improve the life chances of young people aged eight to 25 years in the Medway area. Its annual turnover is £2 million. It has 53 paid employees (42 FTEs) and 15 volunteers.

Regional Headquarters

Corporate Headquarters

56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

European Headquarters

Tamesis
The Glanty
Egham
Surrey, TW20 9AW
UNITED KINGDOM
+44 1784 431611

Asia/Pacific Headquarters

Gartner Australasia Pty. Ltd.
Level 9, 141 Walker Street
North Sydney
New South Wales 2060
AUSTRALIA
+61 2 9459 4600

Japan Headquarters

Gartner Japan Ltd.
Atago Green Hills MORI Tower 5F
2-5-1 Atago, Minato-ku
Tokyo 105-6205
JAPAN
+ 81 3 6430 1800

Latin America Headquarters

Gartner do Brazil
Av. das Nações Unidas, 12551
9° andar—World Trade Center
04578-903—São Paulo SP
BRAZIL
+55 11 3443 1509

© 2012 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. or its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. The information contained in this publication has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information and shall have no liability for errors, omissions or inadequacies in such information. This publication consists of the opinions of Gartner's research organization and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice. Although Gartner research may include a discussion of related legal issues, Gartner does not provide legal advice or services and its research should not be construed or used as such. Gartner is a public company, and its shareholders may include firms and funds that have financial interests in entities covered in Gartner research. Gartner's Board of Directors may include senior managers of these firms or funds. Gartner research is produced independently by its research organization without input or influence from these firms, funds or their managers. For further information on the independence and integrity of Gartner research, see "Guiding Principles on Independence and Objectivity" on its website, http://www.gartner.com/technology/about/ombudsman/omb_guide2.jsp.