A Guide to Social Housing Analytics

In partnership with:

HACT

BEEVER AND STRUTHERS

QAHC

A Fresh Approach

i4H is a new insight tool designed to help drive improvements and deliver value for money using data analytics. Beever and struthers, HACT and QAHC have developed a collaborative partnership to offer this service, ensuring that we consistently meet your needs and continually improve to maximise the performance of your organisation

1. Introduction

The Social Housing performance analytics landscape consists of a wide array of capability. With suppliers, HCA, and consultants approaching analytics very differently, gaining a clear perspective on how to deploy a performance analytics-driven capability can be challenging.

We have therefore produced a layman's guide to performance analytics in order to help our clients make sense of these of these new terms.

In simple terms, the social housing performance analytics landscape consists of:
(a) Reporting and software that allow you to retrieve information from a housing/finance database
(b) Data and statistics-driven predictive modeling that enables customer segmentation

Setting aside the latest buzzwords — such as big data, data mining, and customer intelligence — the following list describes the key categories of performance analytics capability used by social housing organisations.

The relative importance and utility of each category are largely dependent on the sophistication level, budget, and size of the association. However, it is a universal maxim that some combination of performance management reporting and customer segmentation is required to effectively manage the organisation. Accordingly, the following types of analytics are generally available and can be utilised by housing organisations.

2. The Top 10 Performance Usages for Analytics

1. Standard Reports - Typically generated on a regular basis, standard reports describe what has happened in a particular area. They answer the questions "What happened?" and "When did it happen?". They are not useful in making long-term decisions. Examples include monthly or quarterly financial reports.

2. Ad Hoc Reports - Generally, ad hoc reports let you ask questions and request a custom report to find the answer to a specific business question. They answer the questions "How many?", "How often?", and "Where?". A custom report that describes a welfare reform campaign performance is an example of this type of report.

3. On-Line Analytical Processing (OLAP) - OLAP allows you to look at multidimensional data from multiple perspectives. OLAP lets you manipulate the data to efficiently perform analytical operations to answers questions of "How many?" broken down by geography and age. An example of this is sorting and exploring data about different types of customers and their rent payment behavior.

4. Alerts or Triggers - With alerts or triggers, you can learn when you have a problem or opportunity and be notified when something similar happens again in the future. Alerts can appear via email, as a flag within the software, or as red dials on a scorecard or dashboard. They answer the questions "When should I react?" and "What actions are needed now?". An example of an alert or trigger would be an email to a finance officer indicating that interest rates on the bond markets have changed**.**

5 Statistical Analysis - Statistical analysis is about using more complex analytics, like frequency models and regression analysis, to discern patterns and associations within the data. They answer the questions "Why is this happening?" and "What opportunities am I missing?". An analysis determining the association of tenant status and bottom line costs would be an example of this.

6. Forecasting - Forecasting is one of the most useful analytical applications, as it enables effective resource and budget allocation. It answers the questions "What if these trends continue?", "How much is needed?", and "When will it be needed?". As an example, RP's can use forecasting to predict how universal credit payments will affect their overall income, enabling budget allocation and strategy refinement.

7. Segmentation or Descriptive Data - Descriptive data uses for example customer attributes to describe customer behavior or classify customers into groups. Generally, it uses historical behavior to classify individuals, enabling future treatment strategies. It answers the questions "What group or classification does this individual belong to?" and "What characteristics does this individual have?". Examples of segmentation or descriptive data include address, age, income, marital status, presence of children in household, and recent credit history.

8. Predictive Modeling - Predictive modeling analyses historical and comparative data, say about customers to predict a future behavior. It answers the questions "What will happen next?" and "How will it affect my organisation?". Examples of predictive modeling include likelihood to respond to a direct mail solicitation, or the likelihood of not paying rent.

9. Decision Support System (DSS) or Prescriptive Analytics - Prescriptive analytics synthesises data to make predictions and then suggests options to take advantage of the prediction. It describes what you should do and prompts a specific action.

10. Optimisation - Optimisation supports innovation. It takes your resources and needs into consideration and helps you find the best possible way to accomplish your goals, answering the questions "How do we do things better?" and "What is the best decision for a complex problem?". An example using optimisation would be: Given business priorities, budget constraints and available technology, what is the best way to optimise our development spend to meet our development objective?

3. What is Data Mining, Big Data, and Business Intelligence?

The concept of analytics is not new. However, with the proliferation of applications with high volume transactional data and relatively inexpensive infrastructure to store and manage this data, more not for profit sectors are harnessing more data for analytics purposes. This has led to an increasingly sophisticated means of managing and exploiting information.

Generally speaking, the needs of associations are relatively modest compared to other sectors with greater volumes of data and more ample budgets. The following terms are often mentioned by suppliers, agencies and consultants — and more often than not, they are being applied to far more simple concepts than their respective technical definitions.

1. **Data Mining**
   What it is — Data mining is the computational process of discovering patterns in large data sets involving artificial intelligence, machine learning, statistics, and database systems.
   How we misuse the term — Data mining is often misused to mean any form of large-scale data or information processing, for example: collection, extraction, warehousing, analysis, and statistics.
   What we typically mean — Often, we mean the more general term "data analysis" or "analytics."
   Implication for RP's - Generally, this will mean you will require predictive analytics generated by traditional statistical analysis. A statistician will typically deploy software that uses logistic regression or CHAID (goodness of fit) to create a predictive model, score/rating, or segmentation strategy for you. Realistically, sophisticated data mining may not be relevant on your typical size data set.

2. **Big Data**
   What it is — Big data is a collection of data sets so large and complex that it is too difficult to process using standard database management tools or traditional data processing applications. Big data typically requires parallel processing or supercomputing platforms consisting of thousands of servers.
   How we misuse the term — Big data is often misused to mean any large data set that aggregates data from multiple sources.
   What we typically mean —Often the more general terms "large data set" or "multi-sourced database" are more accurate. Implication for associations — fortunately, you may not have big data. Instead, you have "a lot of data."
   Implication for RP's - Generally, if you have a significant amount of data, it can be effectively managed in a standard, cost effective Oracle® / SQL/ Excel database, and accessed via standard query tools systems.

3. **Business Intelligence Application**
   What it is — Business intelligence applications are a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information. They are also often referred to as a "decision support system."
   How we misuse the term — Business intelligence applications are often used to refer to an ad hoc query and reporting tool that extracts information from a data mart or data warehouse. Whilst ad hoc query and reporting is technically a business intelligence application, it is the most primitive.

What we typically mean — Often the more general term of "ad hoc query and reporting tool" is the most accurate, using a "business intelligence platform" to refer to a more sophisticated decision support system that incorporates rules-based decision logic and prescribes specific outcomes or actions.
Implication for RP's — Generally, your CRM, finance or HR system comes with ad hoc query and reporting capability. More sophisticated CRM systems for example will often come with decision support capability.

4. Conclusion

Analytics-driven performance insights can be made extraordinarily easy. Too often, complication stems from collecting data and then trying to use analytics when attempting to determine what to do with the information.

Ideally, if you are able to ask the right question, there is an analytical method that will provide you with the answer or a simple feature that will enable you to query your database to obtain the answer.

For most associations, a simple decision support and reporting and predictive analytics capability offer a direct path to answering your most complicated data analysis challenges.

So, how do you get started? Craft a simple but high impact question like, "How does my historical financial performance compare with the best performing associations and what is the predicted rate of improvement of my peers/sector versus my association?

With our **Housing Performance Analytics** database – we provide an overall financial capability score for your requested peers, that enable you to compare your relative historical performance. We are then able to predict the rate of improvement of your peers and the sector overall. You are then able to compare that with both our theoretical prediction for your association and your budgeted improvement.

With this predictive and prescriptive analytics combination, you are well on your way to deploying a cost-effective analytics strategy providing you with greater performance insights giving you improved decision making.

Appendix 1 – Housing Analytics Examples

Regression analysis is about finding relationships in the data. There are two general approaches to regression that you could take – a per organisation basis and a per tenancy/property basis. For example, per organisation is what the HCA has previously published where you would be looking to say – given your mix of stock, mix of tenants, areas you work in, etc., what is your expected level of expenditure from the regression – if you're below that you're 'efficient', if you're above it you're expensive. They go on to say that being expensive isn't necessarily bad, if boards are consciously choosing to be expensive for some valid reason, like differentiating with high service or quality or whatever.

Per property/tenancy takes the same approach at a more granular level – given all we know about the property, tenants, area, etc., what do we expect to spend on various parts of the business. There are a range of algorithms you can use that are effectively regressions – what you are trying to do is predict the expected level of something (expenditure or service usage in our case) given a range other things that we know. Some of those may give more accurate results than the linear regression that is familiar to most people. Once you have a predicted value for each property/tenancy, you can compare it to the actual value and identify whether there are certain properties that are higher than expected.

This sort of modelling could work within an RP, but the problem would be the variability. You will always have a lot of essentially random factors that will swamp anything that you are interested in at a property level. If tenant A's boiler breaks and needs replacing in a given year, his/her repairs are bound to be much higher than the average that the model predicted for her. The way to make this work is to start grouping properties together in different ways – then if you find that all of your two bed properties are more expensive than they should be, e.g., you might infer that there's something about how you deal with two beds that merits looking at.
The slight downside is, if you did the regression within a single organisation it would almost certainly cancel those sorts of things out. You would put the number of bedrooms into the model, and it would increase its predictions for two bed homes. To make this sort of regression meaningful then you would need to do it for multiple organisations.

Propensity score matching (PSM) this technique will identify with more confidence a relationship between employment status (for example) and repairs costs. It can be used for detecting differences between particular categories of tenants. There are some practical requirements to be able to do a PSM analysis – like you need an accurate record of which category the tenants are in. That is why employment status, where Housing Benefit status gives a very good proxy for employment. There are also some technical requirements – in particular one called common support, which sort of means that the people in the two categories cannot be too dissimilar.

Contact Details

For more information on i4H, or for any queries, our contact details are below.

www.i4housing.org

01926 885 653

info@intel4Housing.com

@i4_housing