

Marketing Research of Construction Sites based on ABC-XYZ Analysis and Relational Data

Aleksandr Konikov^{1,*} and *Gregory Konikov*¹

¹Moscow State University of Civil Engineering , Yaroslavskoe shosse, 26, Moscow, 129337, Russia

Abstract: ABC-XYZ analysis is well known in marketing. It allows identifying sites that yield maximum profits when sold, sites that enjoy stable demand, or sites have both qualities specified above. However, the methods are quite abstract and are not designed to study specific factors that impact the results of ABC-XYZ analysis. Meanwhile, for some applications, particularly for marketing research of construction sites, it is critical not only to identify high-profit and stable sites but also to find out what combination of technical parameters, factors related to their location, transport accessibility, etc. are typical of them. This work suggests an approach to address the issue.

1 Introduction

In marketing and logistics, ABC-XYZ-based research methods are well known for quite a long time [1-9]. This article considers feasibility of using these methods with respect to construction sites. The consideration uses the elements of the set theory providing an opportunity to use a wider approach to the issue, visualize and make most of the findings more intuitive. If compared to the known works, the new feature is that ABC-XYZ analysis is combined with analysis that leverages the use of databases.

This allows not only identifying high-profit and stable sites but also discovering what combination of technical parameters, factors related to their location, transport accessibility, etc. are typical of them.

2 The basic idea of research

The suggested approach is illustrated by the following figure.

* Corresponding author: a.konikov@gmail.com

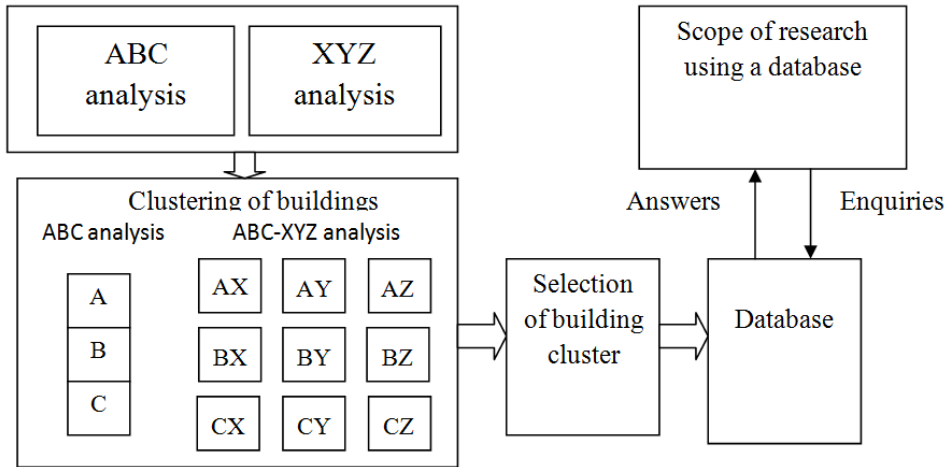


Fig. 1. General scheme of research approach

3 Study on ABC-XYZ analysis using set theory

A departure point of the research is ABC and XYZ analysis.

In the context of the issue to be addressed, ABC analysis could be interpreted as follows: all construction sites are broken down into three categories: A brings up to 80% of the total profits, C returns less than 5%, while the sites of Category B are somewhere in between. For the research purposes, percentage ratios and the number of categories can be varied. XYZ ranges the construction sites based on the stability of profit generation. Stability is suggested to be measured by a value of v , i.e. the ratio of standard deviation from an average value:

$$v = \frac{\bar{x}}{\sigma} * 100\% \quad (1)$$

If the value of v does not exceed 10%, the site generates stable profits and joins group X. If the value of v is between 10% and 25%, the site has less stable profits and joins group Y. If the value of v exceeds 25%, the site has unstable profits and joins group Z.

It should be noted that the results are calculated using Microsoft Excel application. Thus, the analysis can be done with standard functionality of Excel, with no need to involve any third parties. This allows not only avoiding costs but also making the analysis flexible enough, with dynamic changing of analysis conditions and ratio test of the results.

Let's express all construction objects as the set of $P \{p_1, p_2, \dots, p_n\}$, where each element of the set stands for a particular site, i.e.

$$P \{p_1, p_2, \dots, p_n\}; p_i \in P \quad (1)$$

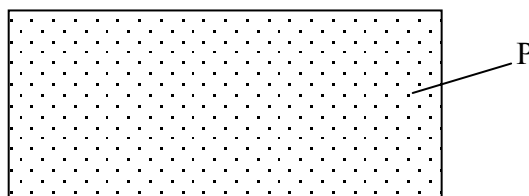


Fig. 2. Set of all sites $P \{p_1, p_2, \dots, p_n\}$

Each element has its own unique ID.

Mathematically, the P set is a combination of three specified subsets:

$$P = A \cup B \cup C \tag{2}$$

The default assumption is that the P set of sites, for which ABC analysis is run, and the set of all construction sites (let it be $W = \{w_1, w_2, \dots, w_\mu\}$) match each other. In general, this won't be the case as illustrated by the following figure:

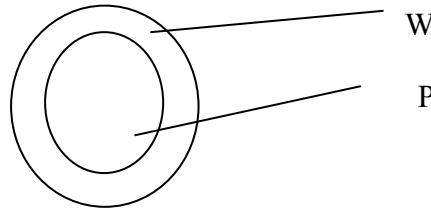


Fig. 3. Set of all construction sites W and set P, which contains sites subject to ABC analysis.

Mathematically, it could be set forth as follows:

$$P \subseteq W \tag{3}$$

$$\forall p \in W \exists w \notin P$$

The physical meaning can be explained as follows. In fact, ABC analysis is run not for all items in the full construction site list. The reasons here can vary. For example, some sites have special status and it's not reasonable to include them in the general ranging list.

In terms of set-theoretical approach, P is the difference of set W and set N of sites put beyond the scope of the research.

$$P = W \setminus N \tag{4}$$

Diagrammatically, it can be illustrated as follows:

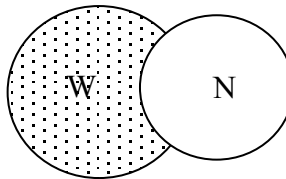


Fig. 4. Difference of W and N sets

So, the construction sites are broken down into three categories: A brings up to 80% of tax proceeds, the total share of C is less than 5%, while the sites of Category B are somewhere in between.

Next, let's take construction sites, for which there are statistical data available. These fields form subset $R = \{r_1, r_2, \dots, r_l\}$, $r_i \in R$

Let's run XYZ analysis to have the set of all sites $R = \{r_1, r_2, \dots, r_l\}$ broken down into three subsets - X, Y, Z.

If taken set-theoretically, set R is a combination of three specified subsets:

$$R = X \cup Y \cup Z \tag{5}$$

Subset X includes construction sites with high stability of demand $v \leq 10\%$ (see above), subset B includes less stable demand sites $10\% > v \leq 25\%$, category C has less stable sites $v > 25\%$

Similarly to ABC analysis, subset $R\{r_1, r_2, \dots, r_l\}$ is a part of the set of all regional sites $W\{w_1, w_2, \dots, w_\mu\}$.

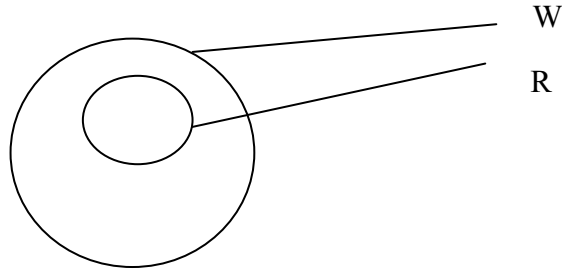


Fig. 5. Set of all construction sites W and set R, which contains sites subject to XYZ analysis.

If put mathematically:

$$R \subseteq W \tag{6}$$

$$\forall r \in W \exists w \notin P$$

It should be noted that the number of screened out sites for XYZ analysis is far greater than for ABC analysis, since beside the factor mentioned above for ABC analysis, the reason for screening out here could simply be the lack of statistics over past periods for a site. Mathematically, set R is the difference of set W and a set of sites put beyond the research scope, let's call it L

$$P = W \setminus L \tag{7}$$

It appears to be a good idea to range sites by two parameters at a time – profitability and stability. If taken set-theoretically, this Q set is an intersection of sets

$$Q = P \cap R = \{q \mid q \in P \wedge q \in R\} \tag{8}$$

The above can be illustrated by the figure:

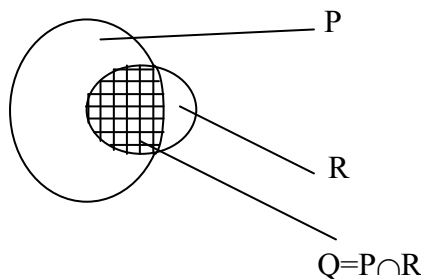


Fig. 6. Set Q as intersection of sets P and R

For set Q, the construction sites can be placed in table ABC-XYZ (Fig. 4)

Table 1. ABC-XYZ table

AX	BX	CX
AY	BY	CY
AZ	BZ	CZ

Where AX cell contains sites that have both high profitability and high stability, AY cell includes sites that have high profitability and medium stability, and so on, down to CZ cell to include sites with low profitability and stability.

4 The next step of the study

The next step, as per scheme 1, is to select a cluster of buildings (construction sites) to be researched in more detail. The term 'cluster' has several interpretations. In this context, a cluster means a group of construction sites selected for the research. Selection depends on particular conditions: say, if there is no statistics required to run XYZ analysis, then it makes sense to choose Group A set (ABC analysis), which brings 80% of the total profits, as a cluster (for research purposes, the value, 80, can be varied). If the required statistics for XYZ analysis is available, then for a cluster we can take the sites of sets AX, AY, BX that are good enough in both reliability and stability.

The sites of the selected cluster are researched using a relational database and relevant business intelligence tools designed to handle databases. A call for the use of databases has the following reasons. It is known that each construction site is followed by detailed documentation: building's engineering certificate, explanatory note, etc. They contain the details of building characteristics, its location, transport accessibility, community services, etc. This information must be added to the database. Moreover, the database can also be enriched with many informal aspects – for example, availability of waters suitable for swimming, etc. An advantage of such relational database is that you can store a huge amount of heterogeneous data in the form, which allows easy processing. A critical point is that such database allows researchers to derive intelligence by processing answers to queries. This should allow identifying factors that impact profitability and sales stability of a construction site. To get better efficiency, it is reasonable to use special analytics tools designed to work with databases (some promising trends are described below).

5 Ways to improve

A development path for this approach could be a replacement of relational database with a more advanced solution - Data Warehouse (DW) and relevant business intelligence technologies: OLAP and Data mining. OLAP is a technology that allows obtaining quick analytical decisions based on multi-dimensional data cube. With Data mining, you can discover hidden regularities by analyzing a large number of heterogeneous data and using a set of smart data analysis technologies. A natural follow-up will be a use of Big Data. We believe that currently this is the most promising technology for processing of large volumes of data. Big Data is extremely suitable for research of large number of construction sites.

Material on new technologies: Data Mining, Data Warehouse, Big Data should occupy enough space in University courses [10].

6 Conclusions

Using the suggested approach, we can identify a group of construction sites that bring the highest profits, a group with most stable profits, and a group of sites, which enjoy both characteristics. Significantly, we can determine which combination of building parameters, factors related to its location, transport accessibility, etc. are typical of each of the groups. The ultimate goal is to help managers and designers of construction sites choose effective solution approaches.

References

1. A. Konikov, G. Konikov. *Procedia Engineering*, **165**, 1052-1056 (2016)
2. A. Konikov, G. Konikov. *AMM*, **584-586**, 2171-2174 (2014)
3. A.I. Konikov, G.A. Konikov, *Int. J. of Applied Engineering Research*, **21(10)**, 42412-42414 (2015)
4. A. I. Konikov, G. A. Konikov. *Logistics and Supply Chain Management*. **5 (64)**, 70-73 (2014)
5. A.I. Konikov, G.A. Konikov, *Improving Efficiency of ABC Analysis with Data Rating, Postgraduate and Candidate*, **2(92)**, 18-20 (2016)
6. A.I. Konikov, G.A. Konikov, *Economics and Business*, **4-1(69-1)**, 727 -730 (2016)
7. A.I. Konikov, T.V. Kuznetchova, G.A. Konikov , *Scientific Works of the Free Economic Society of Russia*, **186**, 269-273 (2014)
8. B.V. Evteev, *SWorld*, **4(41)**, 63-68 (2015)
9. B.V. Evteev *SWorld*, **1(20)**, 68-72 (2014)
10. A.I. Konikov, *Vestnik of the Plekhanov Russian University of Economics*, **4(58)** , 42-46 (2013)