Problems of modeling trade turnover in the retail sector

Mazurenko Mykyta
ORCID ID 0009-0005-1351-8743
Superviser: Ivan Fartushnyi, docent
Igor Sikorsky Kyiv Polytechnic Institute, Ukraine
Superviser: Yuriy Pasenchenko, docent
Kruty Heroes Military Institute of Telecommunications and Information Technology

With the emergence of the first universal methods of statistical modeling, the task of forecasting sales of enterprises in various fields of operation has become widespread. This is done for the purpose of further planning the functioning of the enterprise, determining the standards that play a role in other management systems, such as employee incentives. Also, modeling of such a plan can be used to select optimal values to increase the efficiency of the enterprise.

The retail sector is often characterized by the orientation of sellers and manufacturers to interact with large groups of buyers, which are characterized by various economic, psychological and behavioral factors. Accordingly, the economic results of a firm are not only based on the structure and technology of its operation, but also on the decisions of numerous consumers.

From a mathematical point of view, we can represent turnover in the most detailed way as the sum of the values of all purchases made. Let us present in formula (1) the most detailed formalized representation of a firm's turnover:

$$TO = \sum_{i \in Tr} r(i) \quad (1)$$

where, TO – the company's turnover for the period under study; Tr – the set of purchases made during the study period; i – individual transaction period under study; $r(i)$ – some function that determines the amount of each individual transaction.

Accordingly, the turnover is represented as the sum of revenue received as a result of purchases made by customers. This can be considered the most general representation of turnover, since the set of purchases will be represented as a Cartesian product of an arbitrary number of sets representing time, place, customer, and product category. Accordingly, if properly defined, each such set will uniquely identify a purchase made.

However, such a representation of goods turnover in the case of real systems is practically impossible. A model that would describe each individual transaction cannot be represented analytically. In the case of statistical models, the uncertainty of the number of purchases, the time of their implementation, and the size of the average check remained unknown and difficult to identify. Therefore, for modeling at this level of detail, the only acceptable option is to use a simulation modeling approach, using individual components that determine the purchase in the context of formulating simple rules for agents' actions. Such an approach allows to represent even psychological and behavioral factors and take into account aspects of social dynamics in the decision-making process of customers [1, 2].

In practice, models are often used for aggregated values. For example, in the context of forecasting turnover, the aggregation base can be individual physical locations, brands, certain product categories, or individual items. In practice, retail analysts use the concept of SCU (stock-keeping unit) [3]. Accordingly, by using aggregated data, we can significantly increase the sample on which a model will be implemented or evaluated. Let us present in formula (2) the general form of the function for estimating the conditional indicator of turnover:

$$TO = M_X(x), \quad (2)$$

де M_X – some function representing the turnover estimate; X – historical sample used to parameterize the function; x – observation for which turnover is estimated.

Accordingly, depending on the method and the result of model identification, the resulting characteristic will serve as an estimate of the possible turnover. The main problem with this approach is the high dependence on the quantity and quality of available data. These models require

significantly less labor costs for the identification of its form, analytical analysis and solution, theoretical work laid down in the process of determining the main hypotheses and directly for the construction of the corresponding software product, especially in the case of simulation models. The functions themselves can be used from correlation and regression, factor, variance, cluster analyses or represented by means of data mining. It is also possible to synthesize such estimates with classical models of the subject area, such as the Baas model [4].

The problem with this approach is the complex nature of the indicator itself. In the case of turnover, any purchase can be characterized by the number of purchased items and their cost. Therefore, the next step may be to decompose and build separate models for each selected component. Let us present in formula (3) a variant of the component formula of turnover.

$$TO = E(N \times Cv \times \bar{p}), \tag{3}$$

where N – total number of people who visited the store; Cv – conversion of completed purchases from the total number of visitors; $\bar{p}$ – cost of the average check.

The total turnover is defined as the number of purchases, or the product of the first two components, times the average check. This approach makes it possible to separate analysis and modeling and to identify components that are easier to analyze or have less variation.

The specific choice of one or the other approach depends on: available resources, time allocated for analytics, and the level of information availability for each level of detail.

References:

1. Ben Said, Lamjed & Bouron, Thierry & Drogoul, Alexis. (2002). Agent-based interaction analysis of consumer behavior. 184-190. https://dl.acm.org/doi/abs/10.1145/544741.544787
2. N. Mazurenko and I. Fartushnyi, «Studying Cyclicity in Small Economic System Using Imitational Modeling,» 2022 IEEE 3rd International Conference on System Analysis & Intelligent Computing (SAIC), 2022, pp. 1-5. https://ieeexplore.ieee.org/document/9922980
3. Fader, P. S., & Bruce G. S. Hardie. (1996). Modeling Consumer Choice among SKUs. Journal of Marketing Research, 33(4), 442–452. https://doi.org/10.2307/3152215
4. Fan, Z. P., Che, Y. J., & Chen, Z. Y. (2017). Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. Journal of business research, 74, 90-100.
   https://www.sciencedirect.com/science/article/abs/pii/S0148296317300231