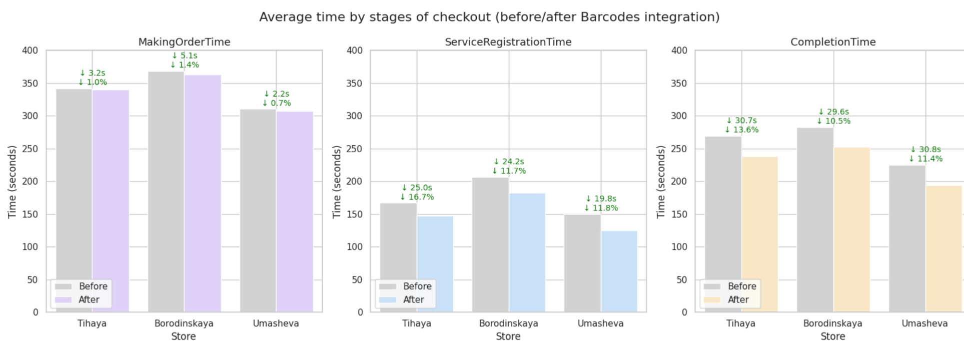
Report results

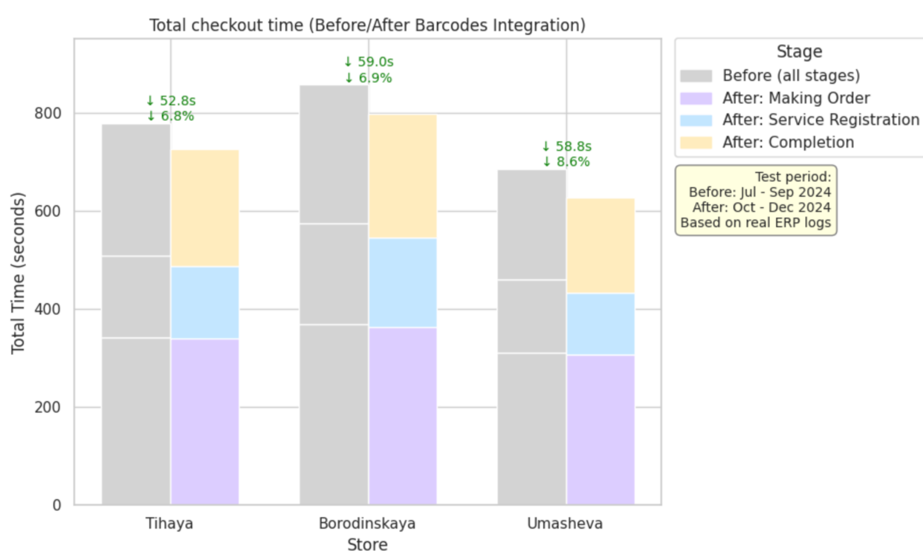
RQ1 - Barcode mechanism results

In order to optimize the customer's order processing time, the ERP program configuration components and the production of barcode cards were developed. The adaptation of the new business process of the customer's path was carried out. The evaluation of the result was carried out in three flagship stores of the company, which have different building characteristics and limitations. For the analysis, data was taken for three months before the implementation and three months after.

Data analysis of the 59,491 observations showed the following changes for the three months before implementation (Q3: July–September) consisting of 24,841 observations and the three months after (Q4: October–December), consisting of 29,650 observations, that make up Customer checkout time.



As the results of the implementation demonstrate, the barcode mechanism has significantly reduced Service Registration Time from 11.7% to 16.7% (from 19.8 s to 25.0 s) and Completion Time from 10.5% to 13.6% (from 29.6 s to 30.8 s). The Making Order Time reflects the time of the client's first contact, which demonstrates minor changes.



The overall decrease in Customer Checkout Time was from 6.8% to 8.6%, which means that each customer spends an average of 52.8 seconds to 59 seconds less, depending on the store.

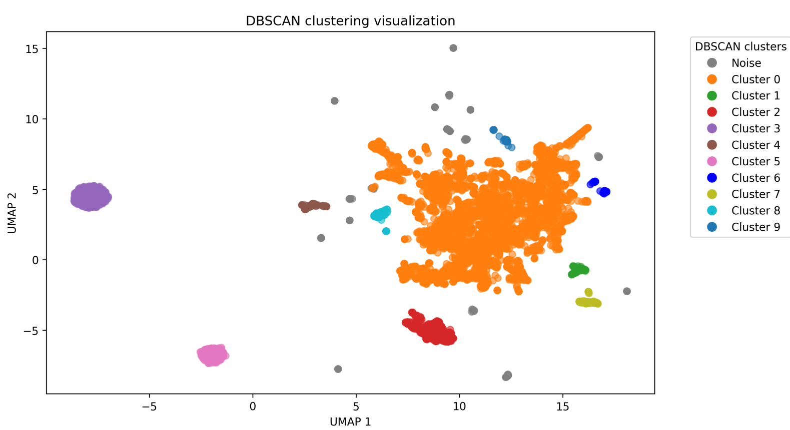
Also, the introduction of barcode cards made it possible to remove printers from the workplaces of sales consultants, which reduced the cost of printing, buying cartridges and paper.

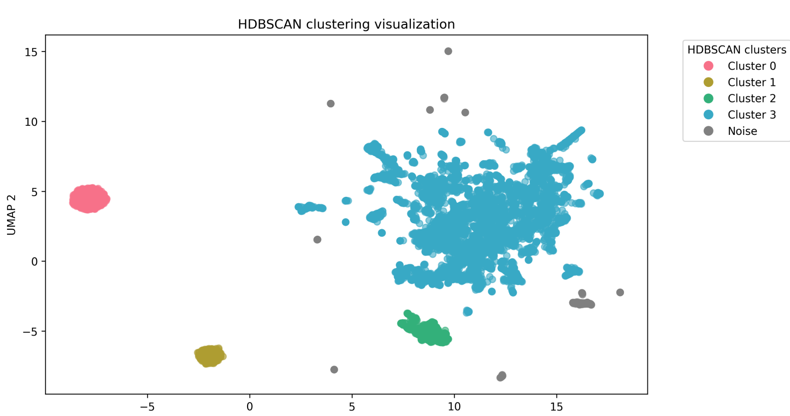
Expert assessment. After the introduction and analytical evaluation of the barcode cards, the company's Executive Director conducted an audit of the integration, which resulted in an overall operating efficiency of 78,050,420 rubles per year across the entire construction retail network. Based on the time elapsed after integration, the efficiency plan is being implemented.

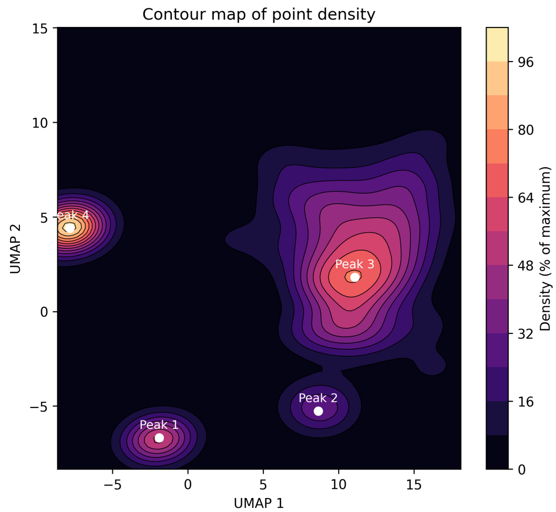
RQ2 – Recommendation System Results

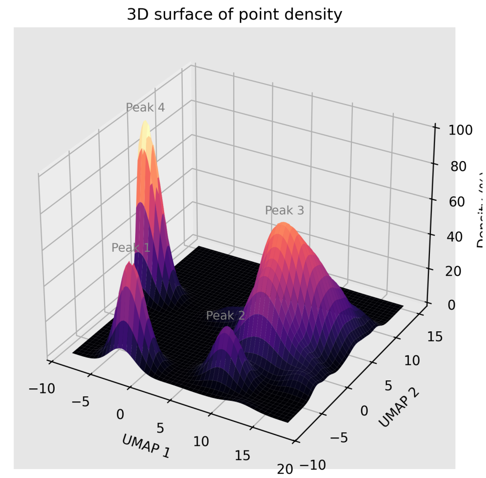
Learning Representations

Before training the models, preliminary clustering of product embeddings was conducted using UMAP and DBSCAN/HDBSCAN algorithms. DBSCAN with ε = 0.87 identified 10 clusters and 243 noise points (approximately 1.9% of the total). Density contour and 3D visualization revealed four prominent peaks, which were used as input for HDBSCAN. The HDBSCAN algorithm identified 4 stable clusters and 266 noise points (around 2%), enabling the integration of cluster labels into the recommendation ranking logic.









Models Evaluation

Model training was performed on a dataset of 492,999 customer orders and 12,881 unique products. Several models were evaluated, including the baseline Global Top recommender, collaborative filtering methods (ALS, Item2Item), and the transformer-based BERT4Rec.

The best-performing model was **BERT4Rec enhanced with HDBSCAN clustering**, achieving:

- **nDCG@20 = 0.2473**
- **Recall@20 = 0.3790**

ALS and Item2Item performed better when trained on binary user-item matrices compared to quantity-based ones. Incorporating MiniLM and SBERT embeddings did not yield significant improvement, likely due to inconsistent naming conventions across product categories.

TABLE 4.1
Performance comparison of recommendation models (nDCG@20 and Recall@20 metrics)

Model	nDCG@20	Recall@20
BERT4Rec + HDBSCAN [27]	**0.2473**	**0.3790**
BERT4Rec + DBSCAN [27]	0.2432	0.3795
Item2Item + Binary Matrix [21]	0.1754	0.2481
ALS + Binary Matrix [21]	0.1225	0.2021
Global Top (Baseline)	0.1146	0.1908
Item2Item + Binary Matrix + MiniLM [21]	0.1086	0.2410
Item2Item + Binary Matrix + SBERT [21]	0.1081	0.2378
Item2Item + Binary Matrix + MiniLM + HDBSCAN [21]	0.1049	0.2307
Item2Item + Binary Matrix + MiniLM + DBSCAN [21]	0.0936	0.1991
ALS + Quantity Matrix [21]	0.0877	0.1442
ALS + Binary Matrix + MiniLM [21]	0.0852	0.1831

Sample Output Examples

Three sample customer baskets were used to compare model outputs. BERT4Rec + HDBSCAN provided the most context-aware and relevant recommendations (e.g., floor protection film for a paint + roller basket, mounting foam for a door frame + screws). In contrast, the baseline model produced generic packaging items that did not reflect the customer's actual intent.

TABLE 4.2
Model comparison based on sample customer baskets

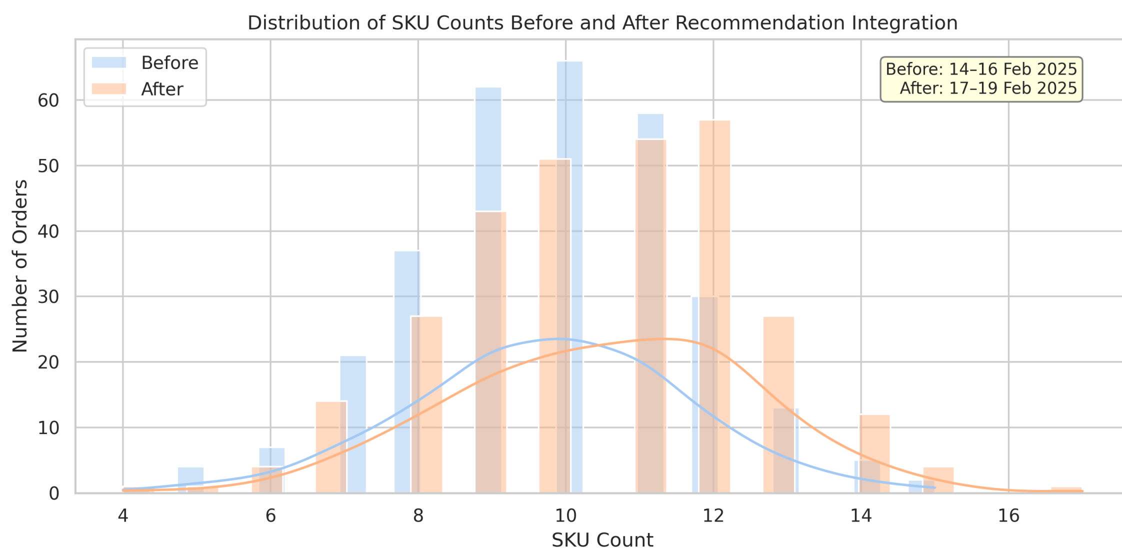
Customer Basket (Input)	Model	Recommended Items
White Interior Paint 10L + Roller for Painting	Global Top (Baseline)	Trash Bags, Plastic Bags, Adhesive Tape
	Item2Item + Binary Matrix	Masking Tape, Stirring Stick, Brush
	BERT4Rec + HDBSCAN	**Masking Tape, Stirring Stick, Floor Protection Film**
Metal Door Frame + Fastening Screws	Global Top (Baseline)	Trash Bags, Plastic Bags, Adhesive Tape
	Item2Item + Binary Matrix	Door Hinges, Fastening Bracket, Insulation Tape
	BERT4Rec + HDBSCAN	**Door Hinges, Mounting Foam, Level**
PPR Pipes + Connectors	Global Top (Baseline)	Trash Bags, Plastic Bags, Adhesive Tape
	Item2Item + Binary Matrix	Valve Kit, Clamps, Sealing Tape
	BERT4Rec + HDBSCAN	**Insulation Sleeves, Valve Kit, Clamps**

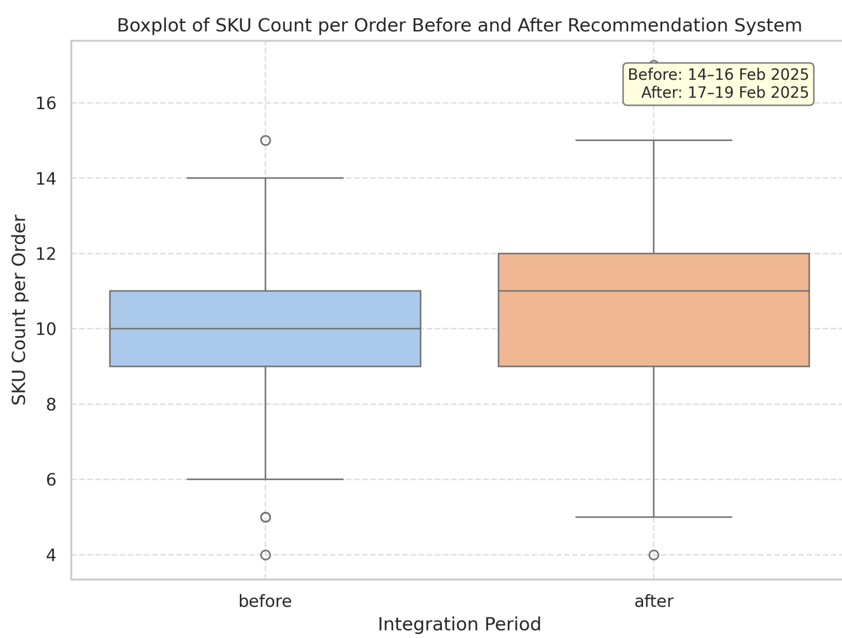
Applied Model Evaluation

For real-world validation, the system was deployed in the "Yumasheva" store and tested across 6 days (3 days before and 3 days after the implementation). Order counts were 306 (pre-deployment) and 296 (post-deployment). After a customer's basket was entered, the recommendation output was shown to the sales consultant on a separate screen.

Results:

- Before deployment: **9.784 SKUs per order**
- After deployment: **10.534 SKUs per order**
- Relative increase: **+0.75 SKUs (+7.660%)**





Expert Assessment

Following the implementation and analytics assessment, the Executive Director of the company conducted an internal audit. It concluded that the recommender system led to an **overall operational efficiency gain of 42,673,540 rubles per year** across the retail network. The improvements were based on results from the three-day pilot and correlated sales data from employee workstations.