Effects of Promotion Mechanisms on Revenue of Sellers in an Online Marketplace

Enver Eren, Mustafa Can Coşkun, Fethi Kahvecioğlu, İbrahim Emre Köksal
Advisors: Taner Bilgiç, Can Özturan
Bogazici University, Computer Engineering Department

Introduction

Our project aims to develop a data-driven discount and promotion strategy tailored to different product and user segments to maximize revenue and sales for JD.com platform.

JD.com, China's largest retailer offered the transactional level data which contains the "full customer experience cycle" that begins with customer "clicks" on product and ends with the delivery to the designated location.

JD.com provides seven tables, which contains 487K order transactions of JD.com and third party sellers occurred in the March 2018. Specific user, product or regional informations are anonymized.

Project Definition and Planning

This project examines the effects of promotion strategies such as direct and quantity discounts on pricing and revenue streams of sellers on JD.com. The examination is conducted through econometric and data-driven descriptive and prescriptive analytics.

Understanding the Data	Descriptive Analytics	Model Building	Scenario Analysis
Exploration of different tables and their summaries. Identifying key relations between each table.	Understanding the table features with numerical and visual analysis. Clustering and feature engineering implementation for simpler structure.	Model implementation on user, product and sales information. Quantifying the relation between promotion and product sales.	Building upon the model, simulating the effect on different promotion strategies on the platform sales for comprehensive advisory.

Data Summary

The data provided by JD.com includes detailed user and SKU identifiers, transactional data, and regional information. Key findings from the data analysis include:

- Orders and Revenue Distribution
- User Demographics
- SKU Attributes

Figure 1: Data Schema

Methodology

- Data Preprocessing: Cleaning and merging the transactional data.
- Regression Tree Construction: Segmenting data based on user and product attributes.
- Simulation Setup: Generating synthetic data for various discount rates.
- Revenue Optimization: Maximizing revenue while considering the cost function.

Regression Tree Structure

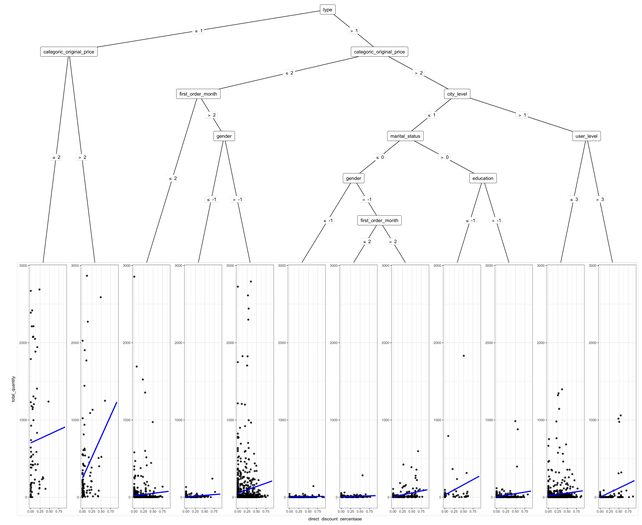

Figure 2: Regression Tree Structure

Simulation Setup

Leaf Node	Regression Equation(Total Quantity)	p-value	MRE	SDRE
1	$703.4 + 224.3 \cdot \text{discount\_percentage}$	0.797	-310.9	780.3
2	$247.6 + 1076.37 \cdot \text{discount\_percentage}$	0.00452	-167.78	523.4
3	$8.941 + 62.723 \cdot \text{discount\_percentage}$	0.0211	-16.94	131.3
4	$611 + 38.212 \cdot \text{discount\_percentage}$	2.38e-06	-3.611	18.51
5	$52.392 + 172.508 \cdot \text{discount\_percentage}$	9.06e-05	-49.39	234.8
6	$2.2508 + 9.7997 \cdot \text{discount\_percentage}$	1.96e-05	-1.251	8.572
7	$5.120 + 16.760 \cdot \text{discount\_percentage}$	0.004909	-3.120	18.5
8	$5.662 + 100.216 \cdot \text{discount\_percentage}$	2.13e-08	-4.66	62.73
9	$30.57 + 263.19 \cdot \text{discount\_percentage}$	0.0559	-27.57	240.9
10	$5.792 + 80.044 \cdot \text{discount\_percentage}$	0.000261	-4.79	74.88
11	$20.361 + 68.033 \cdot \text{discount\_percentage}$	0.00016	-19.36	112.3
12	$1.509 + 233.525 \cdot \text{discount\_percentage}$	5.28e-06	-0.51	130.4

Table 1: Regression Equations, p-values, Mean Residual Errors, and Standard Deviations of Residual Errors

n : Number of SKU's in the leaf

$$\text{Revenue} = \sum_{i=1}^{n} \text{Generated Quantity}_i \cdot \text{Original Unit Price}_i \cdot (1 - \text{Discount Rate}) \quad (1)$$

$$\text{Cost} = \sum_{i=1}^{n} \text{Generated Quantity}_i \cdot \text{Original Unit Price}_i \cdot \text{Discount Rate}$$

Optimization Results

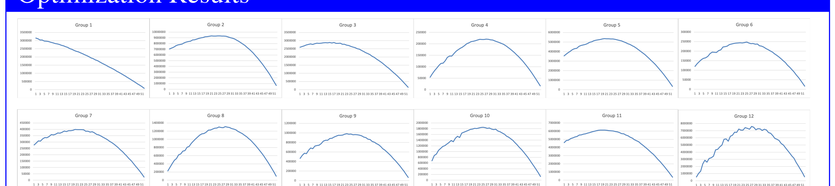

Figure 3: Expected Revenue vs. Discount Rate

Optimization Model

i : Index of leaves

$$x_{i,dr} = \begin{cases} 1 & \text{if discount rate } dr \text{ is assigned to leaf } i \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Maximize} \quad Z = \sum_{i=1}^{n} \sum_{dr \in D} x_{i,dr} \cdot \text{revenue}_{i,dr} \quad (2)$$

$$\text{subject to} \quad \sum_{dr \in D} x_{i,dr} = 1 \qquad \forall i \in \{1, \ldots, n\} \quad (3)$$

$$\sum_{i=1}^{n} \sum_{dr \in D} x_{i,dr} \cdot \text{cost}_{i,dr} \leq \text{Budget} \quad (4)$$

$$x_{i,dr} \in \{0, 1\} \qquad \forall i \in \{1, \ldots, n\}, dr \in D \quad (5)$$

Optimization Results

Optimized Decision / Current Strategy

Group	Discount Rate	Expected Revenue	Cost	Real Discount Rate	Real Revenue	Real Cost
1	0,00	3.156.142,4	0	0,08	4.866.527,74	808.387,44
2	0,12	8.142.720,74	1.110.371,01	0,09	11.056.153,95	2.158.638,8
3	0,06	2.738.589,46	174.803,58	0,08	1.594.367,91	345.253,76
4	0,32	204.829,42	96.390,31	0,06	133.687,81	32.528,34
5	0,20	4.958.762,02	1.239.690,50	0,08	4.371.370,26	992.592,74
6	0,22	221.589,03	62.499,47	0,12	170.544,29	61.003,64
7	0,18	370.721,85	81.377,97	0,12	286.772,50	65.632,16
8	0,26	1.102.665,96	387.423,17	0,18	564.489,29	417.442,50
9	0,20	844.650,20	211.162,55	0,15	452.823,02	253.977,76
10	0,34	1.643.352,99	518.953,58	0,12	672.728,43	320.547,71
11	0,32	5.514.404,27	751.964,22	0,14	4.524.994,46	1.885.953,85
12	0,32	6.775.848,23	3.188.614,46	0,12	605.289,65	482.833,34
Total		35.674.276,6	7.823.270,8		29.299.749,3	7.824.792,1

Figure 4: Constant Budget

Optimized Decision / Current Strategy

Group	Discount Rate	Expected Revenue	Real Discount Rate	Real Revenue	Real Cost
1	0,00	3.156.142,4	0,08	4.866.527,7	808.387,4
2	0,36	9.343.631,2	0,09	11.056.154,0	2.158.638,9
3	0,24	2.882.437,8	0,08	1.594.367,9	345.253,8
4	0,48	2.208.839,8	0,06	133.687,8	32.528,3
5	0,34	5.352.073,2	0,08	4.371.370,3	992.592,7
6	0,44	2.476.595,0	0,12	170.544,3	61.003,6
7	0,34	398.517,1	0,12	286.772,5	65.632,2
8	0,5	1.311.760,0	0,18	564.489,3	417.442,5
9	0,4	9.810.400,8	0,15	452.823,0	253.977,8
10	0,44	1.848.072,3	0,12	672.728,4	320.547,7
11	0,34	6.131.854,6	0,14	4.524.994,5	1.885.953,9
12	0,48	7.608.083,3	0,12	605.289,7	482.833,3
Total		30.482.155,54		29.299.749,31	7.824.792,1

Figure 5: Unlimited Budget

Optimized Decision / Current Strategy

Group	Discount Rate	Expected Revenue	Cost	Real Discount Rate	Real Revenue	Real Cost
1	0,00	3.156.142,4	0	0,08	4.866.527,7	808.387,4
2	0,30	9.177.154,7	3.933.066	0,09	11.056.154,0	2.158.638,9
3	0,14	2.838.651,9	462.106	0,08	1.594.367,9	345.253,8
4	0,40	217.122,4	144.748	0,06	133.687,8	32.528,3
5	0,28	5.226.375,1	2.032.479	0,08	4.371.370,3	992.592,7
6	0,32	240.346,8	113.104	0,12	170.544,3	61.003,6
7	0,34	398.517,1	205.297	0,12	286.772,5	65.632,2
8	0,44	1.303.085,5	1.023.853	0,18	564.489,3	417.442,5
9	0,32	948.281,1	446.250	0,15	452.823,0	253.977,8
10	0,32	1.778.117,5	836.761	0,12	672.728,4	320.547,7
11	0,28	6.116.056,4	2.378.466	0,14	4.524.994,5	1.885.953,9
12	0,36	7.232.514,3	4.068.289	0,12	605.289,7	482.833,3
Total		38.632.365,22	15.644.420,64		29.299.749,31	7.824.792,11

Figure 6: Doubled Budget

Optimized Decision / Current Strategy

Group	Discount Rate	Expected Revenue	Cost	Real Discount Rate	Real Revenue	Real Cost
1	0,00	3.156.142,4	0	0,08	4.866.527,7	808.387,4
2	0,06	7.611.391,2	485.833	0,09	11.056.154,0	2.158.638,9
3	0,00	2.591.948,7	0	0,08	1.594.367,9	345.253,8
4	0,20	167.289,4	41.822	0,06	133.687,8	32.528,3
5	0,12	4.540.525,5	619.163	0,08	4.371.370,3	992.592,7
6	0,14	195.079,4	31.757	0,12	170.544,3	61.003,6
7	0,14	358.222,6	58.315	0,12	286.772,5	65.632,2
8	0,22	1.010.048,0	284.885	0,18	564.489,3	417.442,5
9	0,18	827.349,3	181.613	0,15	452.823,0	253.977,8
10	0,15	1.386.026,4	244.593	0,12	672.728,4	320.547,7
11	0,06	5.087.515,7	324.735	0,14	4.524.994,5	1.885.953,9
12	0,22	5.797.455,2	1.635.180	0,12	605.289,7	482.833,3
Total		32.728.993,96	3.907.897,05		29.299.749,31	7.824.792,11

Figure 7: Halved Budget

Recommendations Based on the Constant Budget Scenario:

- For categories 2, 4, 5, 6, 7, 8, 9, 10, and 12, it is recommended to increase the discount rates to boost revenue.
- For the remaining groups, lowering the discount rates is advised to optimize revenue streams.

Results

The regression tree analysis revealed how different user segments and product categories respond to varying promotions. Simulations indicated significant revenue potential under various promotional scenarios.

Findings, Recommendations, and Assumptions:

- Groups 1 and 2 (Type 1) where JD.com is the reseller can implement strategies more easily due to direct control, allowing flexibility in adjusting discounts and promotions.
- Current budget data is not from JD.com. Using real budget data with actual costs will yield more accurate results, ensuring effective revenue optimization.
- The new strategy is recommended for each leaf node, assuming JD.com has no fulfillment or network capacity constraints, critical for smooth implementation of discount and revenue strategies.

Conclusion

Our study demonstrates that strategic promotional campaigns can significantly enhance revenue and sales on e-commerce platforms. Tailored promotional strategies based on data-driven insights are crucial for optimizing revenue.