

BADM_ Group Project:
Price arbitrage for Mom & Pop Store Product Traders

Group 7B

Abhinav Verma	61310725
Bhaskar Vedula	61310424
Raphael Schilling	61319001
Rhianna Prabhu	61310319
Vivek Aranha	61310554

Business Goal



Main Idea	We are consultants to Mom & Pop Stores, and our goal is to help them make profits by buying products through Bargain.in that are likely to have a price increase within the next 24 hours. These product traders utilize information asymmetry to sell at high prices to their customers.
Stakeholder	Mom & Pop Stores
Business benefit	Price arbitrage
Success measurement	Better prediction than Bayes Naive rule

Datamining Goal



Analytics objective

The data mining goal is to recommend the top 10 brands, across various products, for sale to product traders. These brands are carefully selected through a logistic regression model

Task

Transform input variables (e.g. Brand, day of week, Average Rating) to likely hood of "price up"

Kind of task

Supervised, predictive

Main outcome variables

Price Up – Categorical (1=Price Up, 0 =No Price Up)

Sample Data Used

RowID	Product	Brand Airtime	Brand Arise	Brand Htc	Brand Karbonn	Brand Lg	Brand Micromax	Brand Motorola	Brand Nokia	Brand Samsung	Brand Sony	Av Rating	listPrice	Shipping Period	Day of the week	Price Up
7361	Samsung Galaxy Ace S5830	0	0	0	0	0	0	0	0	1	0	4	10,300	NA	4	0
5227	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,338	NA	3	0
5230	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,489	NA	5	0
5231	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,489	NA	5	0
5236	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,338	NA	5	0
5237	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,338	NA	5	0
5239	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,338	NA	5	0
5244	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,338	NA	6	1
5245	Airtyme Flaunt (Black)	1	0	0	0	0	0	0	0	0	0	4	2,595	NA	6	0

~4,000 rows of Data used after initial processing
Dimension Reduction to top 10 brands

Data-Mining Models

We used a Logistic Regression model to predict likelihood of price increasing within 24 hours

Input Variables: Dummies for Airtym, Arise, HTC, Karbonn, LG, Micromax, Motorola, Nokia, Samsung; Average Rating, Day of Week (Sony Brand used as reference)

Output Variable: Price Up (1 or 0)

Error Rate: 2.23%



Benchmark Model: Naïve Bayes

Input Variables: Dummies for Airtym, Arise, HTC, Karbonn, LG, Micromax, Motorola, Nokia, Samsung; Average Rating, Day of Week (Sony brand used as reference)

Output Variable: Price Up (1 or 0)

Error Rate: 2.31%

Logistic Regression Model outperformed the Benchmark in terms of Error Rate

Key Recommendations

- We recommend Mom & Pop Product Traders to use our logistic regression model to buy products across 10 recommended brands
- Our model takes into account products whose prices will be available within 1 day – This reduces the inventory holding cost of the trader
- Major Risks
 - Change in customer preferences / Products may become obsolete – Risk of over-stocking / liquidation risk
 - Recommended products might be out of stock