



RECOMMENDATION SYSTEM

Business Case developed by:

- Ernesto Aguilar
- Fernanda Zippinotti
- Filipa Cerqueira
- Laura Castro

AGENDA



BUSINESS SITUATION



Background



E-commerce and online sales had rapidly expanded over many markets due to the logistics and time saving convenience.



Online grocery market is not the exception, in India, this market has grown **12%** between 2010 and 2015

About bigbasket

- **Bigbasket.com** was established as online grocery and food store in 2011 in India.
- BigBasket identified an opportunity to increase their sales by addressing the customers with a **recommendation system.**



BUSINESS OBJECTIVES

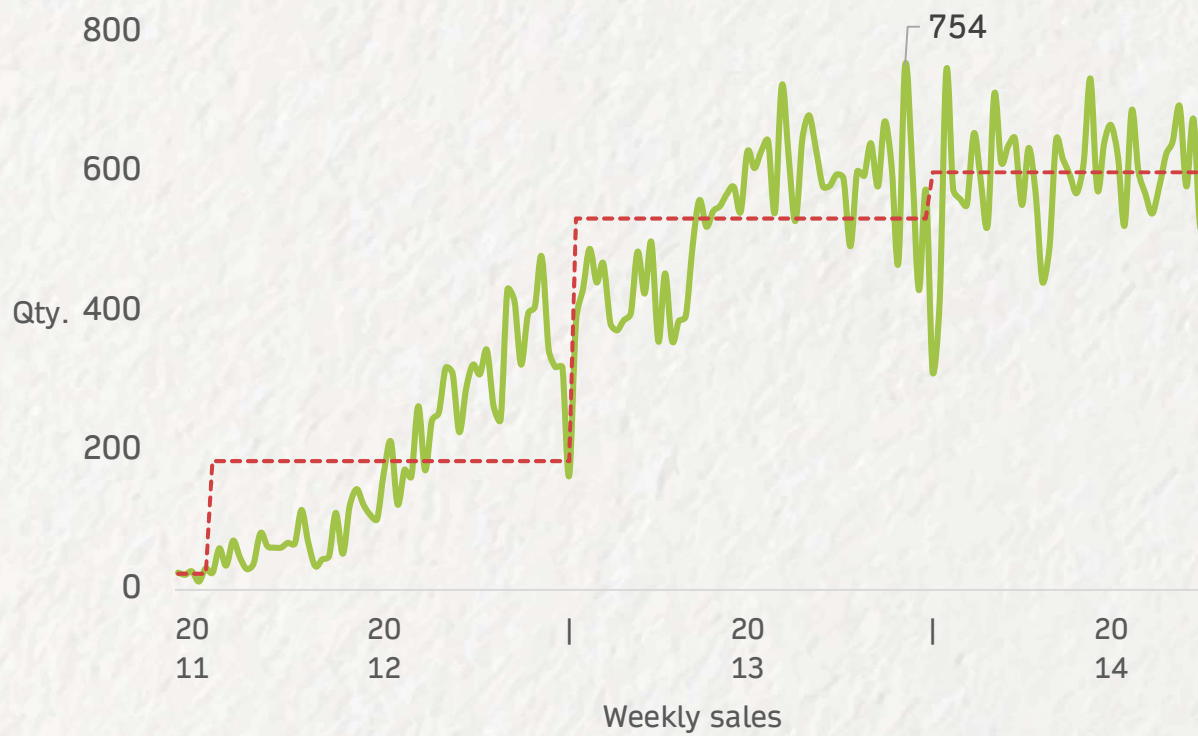
- Better buying experience.
- Increase sales by product recommendation
- Decrease costs associated to second small deliveries

BENEFITS

- Increase the average order value
- Better customers' engagement
- Have a better control on retailing and logistics.



SALES OVER TIME



Avg. Qty. 2012

158

Avg. Qty. 2013

533

Avg. Qty. 2014

600

Max. Qty.

754

CUSTOMER SEGMENTATION

Frequent High Orders Members

18%

- Buy Frequently
- **Average Quantity per Order: 10**
- Recency: 1 Week

Frequent Members

42%

- Buy Frequently
- **Average Quantity per Order: 8**
- Recency: 3 Weeks

40%

Lower Quantities Members

- Most frequent buyers
- **Average Quantity per Order: 6**
- Recency: 3 Week

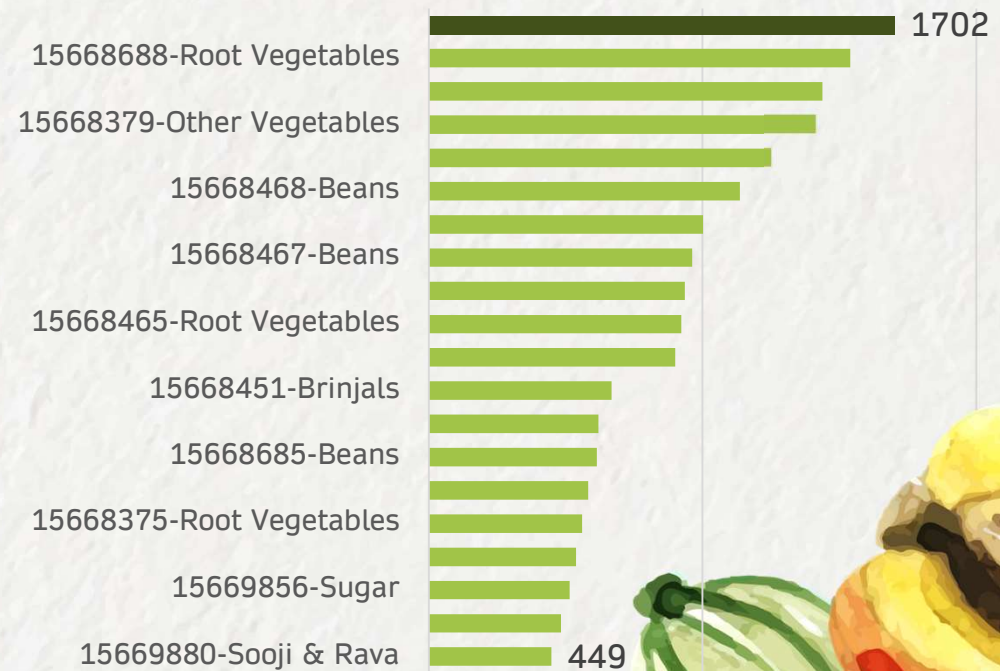


BEST SELLING

Top sold Descriptions

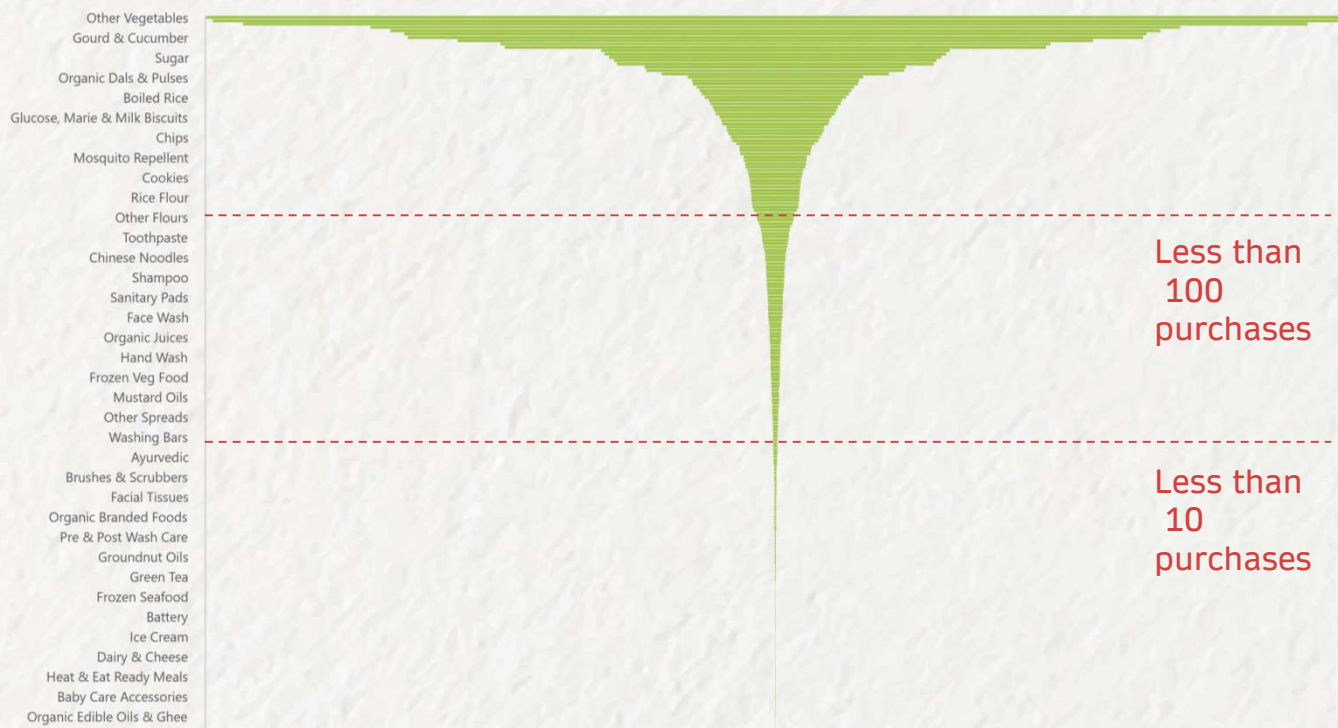


Top 20 products



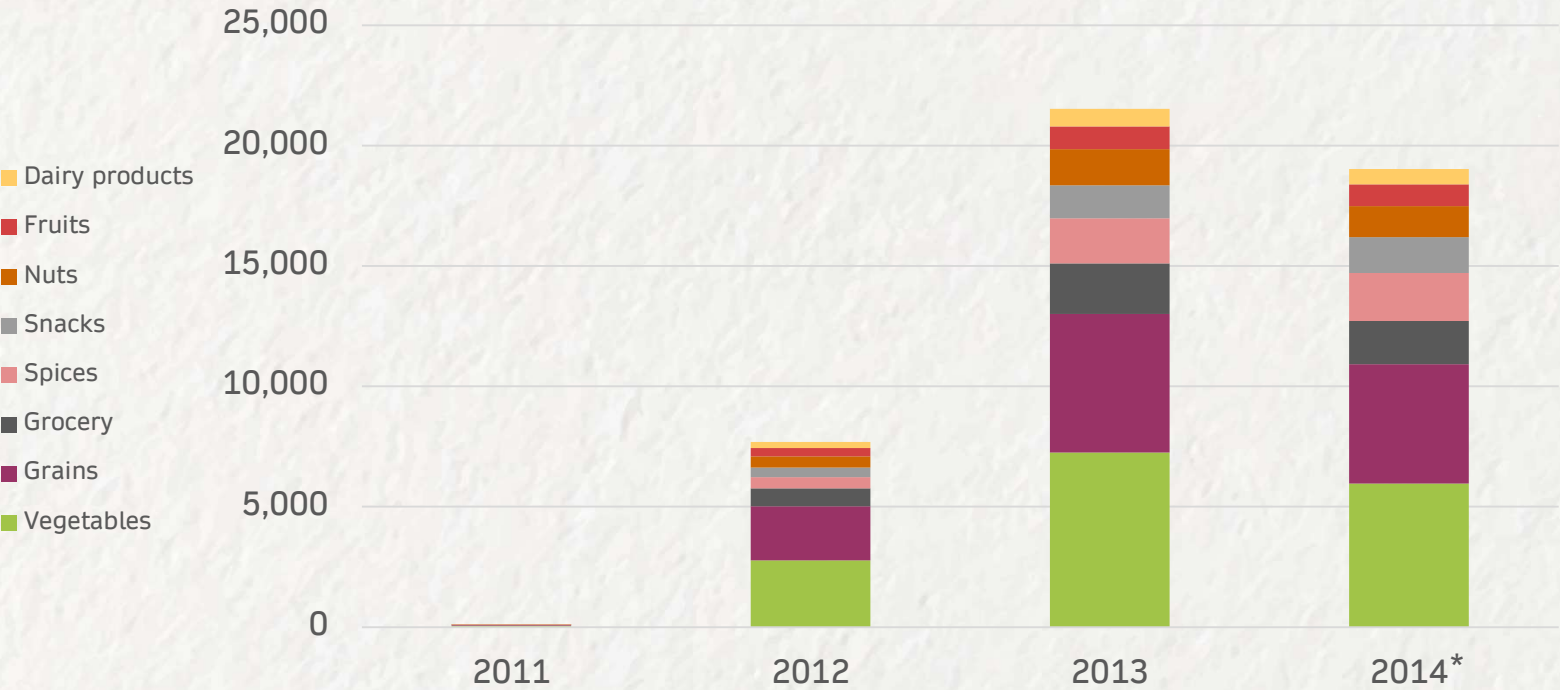
BEST SELLING

Qty. sold by Description



SALES BY CATEGORIES

Top sold Categories



* Year not complete

DATA PREPARATION



concatenation of SKU
code and the category
Description



Same Days Orders
Grouped as One



Orders with less than 5
items excluded

Dealing with the 'Forgetting
Problem'



MODELLING STEPS

Provided Data

Data Preparation

Train

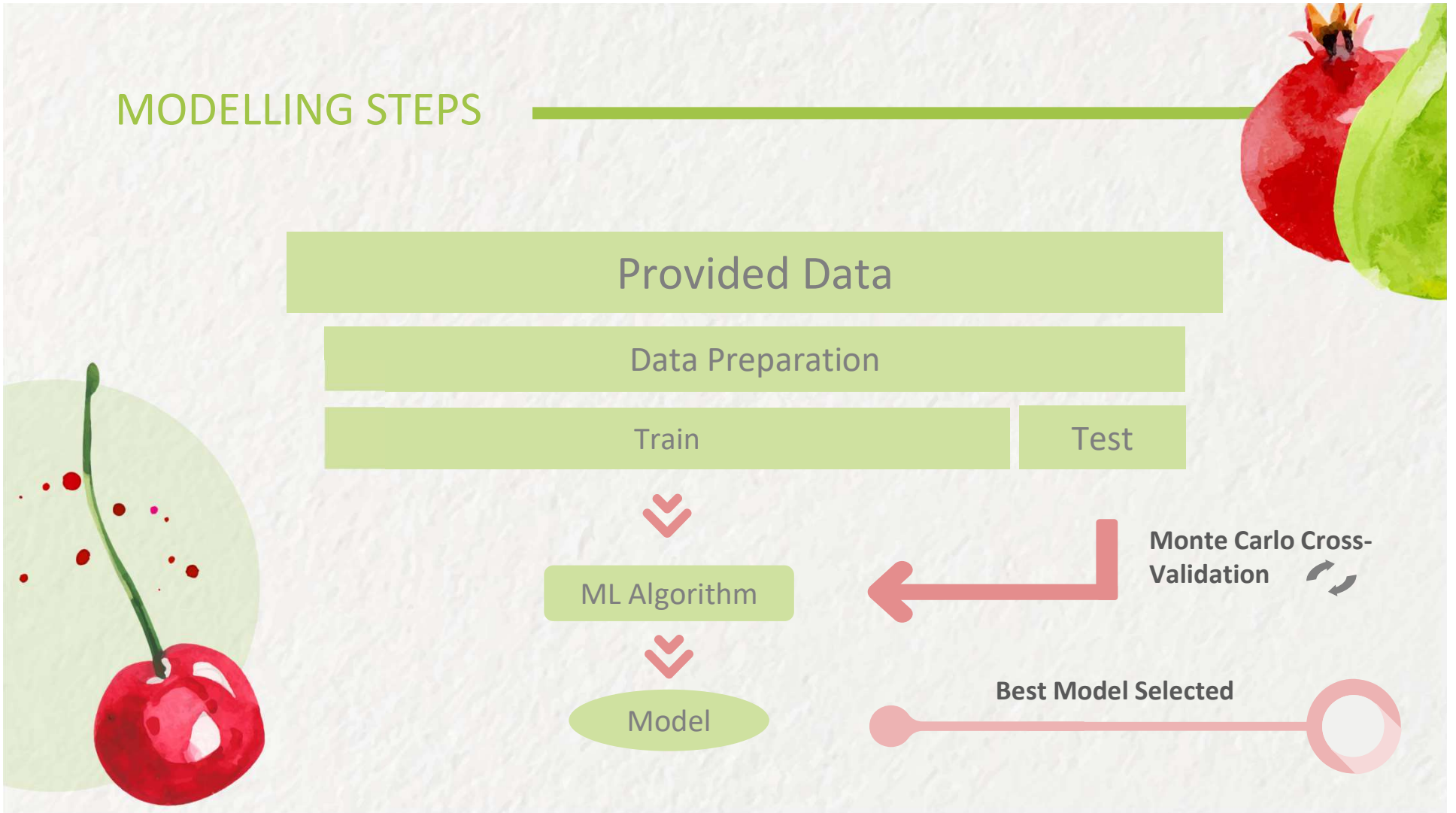
Test

ML Algorithm

Model

Monte Carlo Cross-
Validation

Best Model Selected



MODEL SELECTION

Evaluation Metrics

- Recall
- Precision
- Regular Recommendation
- Regular Shop

Smart Basket Problem

1

Page Rank

2

Most Frequent

3

Hybrid
(Model 1 + 2)

Did You Forget Problem

4

Page Rank

5

Similarity
Recommendation:
Cosine

6

Similarity
Recommendation:
Jaccard

7

Similarity
Recommendation:
Dice

8

Market Basket
Analysis



SMART BASKET PROBLEM



M56xx9

	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	
Beans	5	4	3	4	6	1	2	4	3	-	4	2	38
Ground Coffee	4	3	4	3	2	5	3	2	2	2	-	1	31
Namkeen	4	9	9	-	-	-	-	-	-	-	2	4	28
Whole Spices	7	-	-	2	-	5	-	-	7	6	-	-	27
Sugar	1	-	2	-	5	2	-	5	-	1	2	-	18
Root Vegetables	4	-	-	-	-	2	-	-	-	-	-	2	8
Sooji & Rava	-	-	-	-	2	2	-	-	2	2	-	-	8
Sunflower Oils	-	-	-	-	3	2	-	3	-	-	-	-	8
Toor Dal	-	-	-	-	2	-	1	3	2	-	-	-	8
Yogurt & Lassi	-	-	-	-	-	-	-	2	2	-	-	4	8
Almonds	-	2	1	1	-	-	1	1	-	-	1	-	7
Corn Snacks	-	5	-	-	-	-	-	-	-	-	-	2	7
Cream Biscuits	2	2	3	-	-	-	-	-	-	-	-	-	7
Organic Flours	-	2	1	-	-	-	-	-	-	-	2	1	6
Soup	-	-	2	-	-	-	2	-	-	-	-	2	6
Avalakki / Poha	-	-	-	2	-	-	-	1	-	2	1	-	6
Bread	-	1	1	-	-	-	-	-	1	-	1	1	5
Ice Creams & Desserts	-	-	-	5	-	-	-	-	-	-	-	-	5
Other Dals	-	-	-	-	-	-	2	-	-	2	-	-	4
Other Dry Fruits	-	4	-	-	-	-	-	-	-	-	-	-	4
	27	32	26	17	20	19	11	21	19	15	13	19	Σ

SMART BASKET PROBLEM

- ❑ Select the right basket to the right customer
- ❑ Increase the customer experience



M56xx9

His purchased:

Cashews
Ground Coffee
Toor Dal
Sugar
Sunflower Oils
Beans
Urad Dal

Smart basket recommendation:

Ground Coffee	Other Rice Products
Cashews	Yogurt & Lassi
Toor Dal	Raw Rice
Beans	Avalakki / Poha
Root Vegetables	Sooji & Rava
Moong Dal	Urad Dal
Sunflower Oils	Sugar
Namkeen	Whole Spices
Besan	

SMART BASKET PROBLEM



M56xx9

His purchased:

Cashews
Ground Coffee
Toor Dal
Sugar
Sunflower Oils
Beans
Urad Dal

In the smart basket?

Y
Y
Y
Y
Y
Y
N

Precision: **30%**

Recall: **85%**

SMART BASKET PROBLEM



M56xx9

His purchased:

Cashews
Ground Coffee
Toor Dal
Sugar
Sunflower Oils
Beans
Urad Dal

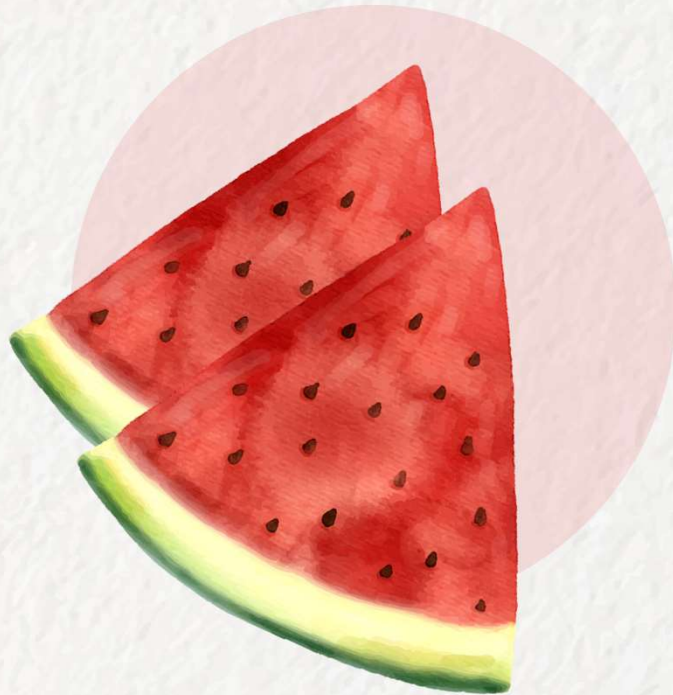
In the smart basket?

Y
Y
Y
Y
Y
Y
N

Recall: **85%**

$$6/7 = 0.85$$

SMART BASKET PROBLEM



Overall Evaluation

	RANDOM	MIX MODEL
Recall	0.132 +/- 0.12	0.436 +/- 0.21

DID YOU FORGET?

- ❑ Suggest products that are related to the ones that the customer purchased
- ❑ Increase the customer experience



M56xx9

His purchased:

Cashews
Ground Coffee
Toor Dal
Urad Dal
Sugar
Sunflower Oils
Beans

Did you forget recommendations:

Sugar
Sunflower Oils
Chips
Sooji Rava
Banana

DID YOU FORGET?



M56xx9

His purchased:

Cashews
Ground Coffee
Toor Dal
Urad Dal
Sugar
Sunflower Oils
Beans

In the smart basket?

-
-
-
-
Y
Y
N

Precision: **40%**

Recall: **66%**

DID YOU FORGET?



M56xx9

His purchased:

Cashews
Ground Coffee
Toor Dal
Urad Dal
Sugar
Sunflower Oils
Beans

In the smart basket?

-
-
-
-
Y
Y
N

Recall: **66%**

$2/3=0.66$

DID YOU FORGET?

Overall Evaluation

	RANDOM	PAGE RANK
Recall	0.024 +/- 0.08	0.2205 +/- 0.217



NEW CUSTOMER



NEW CUSTOMER

Uses information from
all other customers



SMART BASKET PROBLEM

	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	
Other Vegetables	381	334	380	304	416	429	492	464	464	291	312	318	4,585
Beans	368	312	359	271	359	414	443	449	466	361	397	335	4,534
Root Vegetables	205	179	361	326	438	442	439	430	459	287	297	246	4,109
Organic F&V	269	167	190	138	168	196	232	237	157	147	-	210	2,111
Other Dals	159	224	-	123	316	307	45	187	278	42	235	144	2,060
Whole Spices	152	132	159	149	171	162	307	319	140	184	-	137	2,012
Gourd & Cucumber	110	102	239	115	29	32	187	68	72	68	222	90	1,334
Brinjals	102	15	24	-	-	-	-	56	79	67	-	-	343
	1,746	1,465	1,712	1,426	1,897	1,982	2,145	2,210	2,115	1,447	1,463	1,480	Σ

NEW CUSTOMER

Smart basket Evaluation

	RANDOM	PAGE RANK
Recall	0.132 +/- 0.12	0.366 +/- 0.245

Did you forget Evaluation

	RANDOM	COSINE MODEL
Recall	0.024 +/- 0.08	0.191 +/- 0.234



MODEL DEPLOYMENT

Smart Basket Process



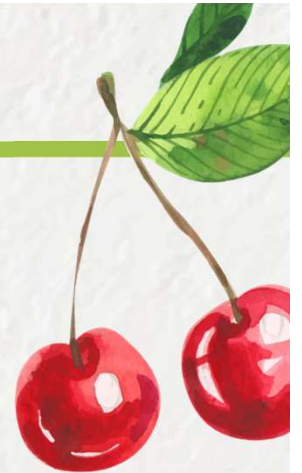
Disposal of customer's Smart Basket



Did You Forget Process



Disposal of Items that customer could have forgot



MAINTENANCE PLANS



PRODUCTS

New products need to be added and products that are no longer available need to be removed.



TIMELINE

Considered only the five most recent years from the database.

CONSIDERATIONS FOR MODEL IMPROVEMENT



PRICE

Price will influence purchasing choice.



DEMOGRAPHIC DATA

More accurate results.

Clustering



STOCK

Products that aren't available shouldn't be suggested.

Conclusions



Customer Satisfaction



**Accurate product
recommendations**



Sales increase



THANKS!



Does anyone have any questions?

CREDITS: This presentation template was created by
Slidesgo, including icons by Flaticon, and infographics
& images by Freepik.

Please keep this slide for attribution.