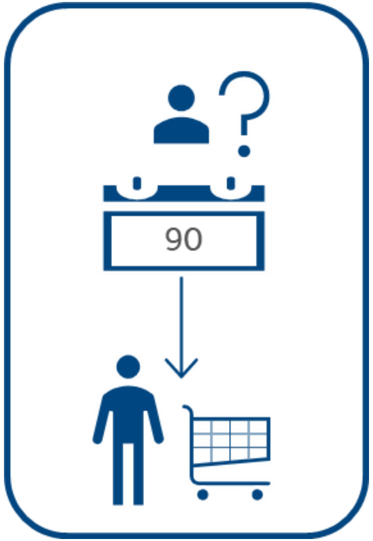
Engagement and Retention Strategy: Recent Shoppers

Tanmay Sinnarkar
8/30/2022

Background & Objective



Background

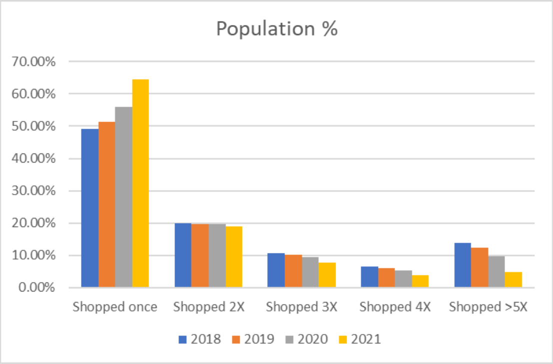
- Majority of the customers are one-time shoppers; the business is looking to contact the customers that we believe will make a purchase and identify the customers that will not come back (One-time shoppers).

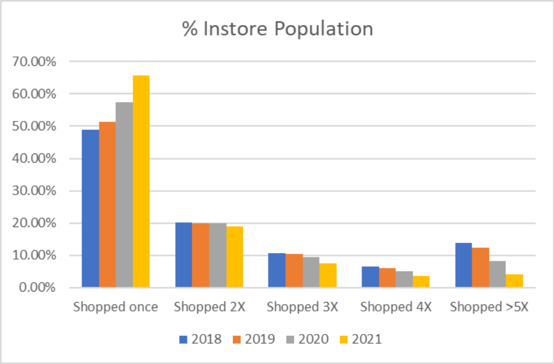
Objective

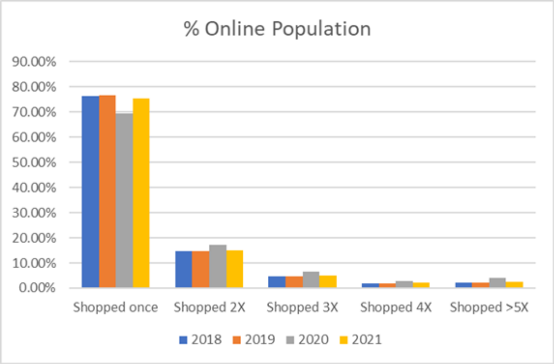
- To predict the likelihood of a customer will make their next purchase 90 days from the day they made their last purchase.
- Creating an opportunity for repeat purchases through personalized Emails.

Understanding Customer Engagement : Shopping Frequency

- A YoY rise could be observed in the % of one-time shoppers with a decrease in the % of returning customers.
- One-time shoppers % higher in 2021 with **64%** compared to **51%** in 2019, repeat shoppers % is lower in 2021 compared to 2019.
- Instore: One-time shopper, **66%** in 2021, higher compared to **51%** in 2019 and repeat shoppers % is lower in 2021 compared to 2019
- Online: One-time shopper, **75%** in 2021, lower compared to **77%** in 2019, repeats shoppers trending higher in 2021 compared to 2019



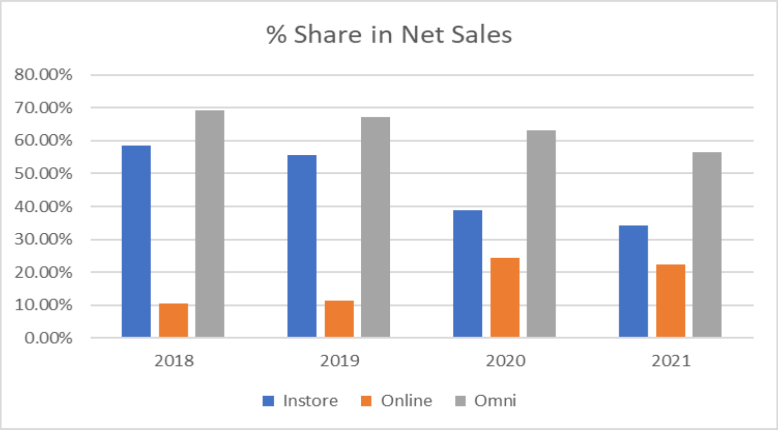




Importance of Consecutive Shoppers

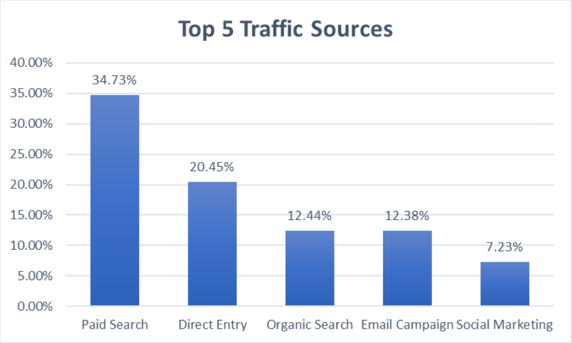
Behavior of the customers who shop at least once consecutively within 90 days in a year:

- At least **30%** of the total customers shop twice within 90 days in a year.
- Of these customers, their total purchase within the year accounts to at least **56%** of the total yearly net sales overall.
- At least 11% of the Total Yearly Net Sales considering Online Shopping.
- At least 34% of the Total Yearly Net Sales considering Instore Shopping.

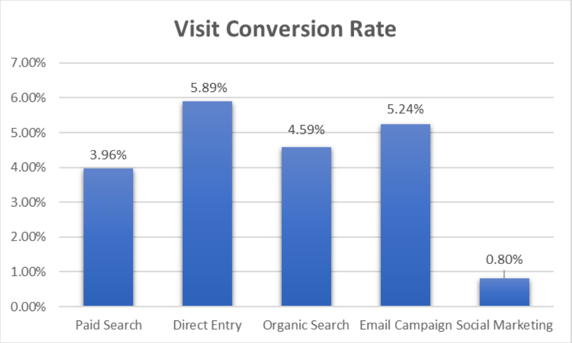


Channel Level Analysis

- Email Campaigns account for 12.38% of total online visits.
- In 2021, 45.84M visits through Email led to a 5.24% conversion rate.
- From 19.2B Emails being sent in 2021, the resulting visit rate is 0.27%.

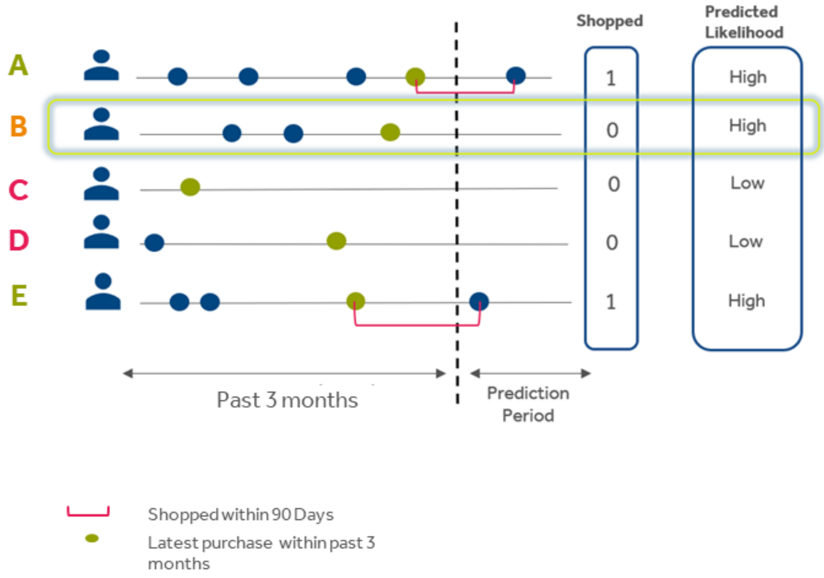


Emails Sent	Visit Count	Visit Rate	Converted Count	Visit Conversion Rate	Product Browse Count	Abandoned Browsed Product Count	Product Abandonment Rate
19.24B	45.84M	0.24%	2.40M	5.24%	2.01B	2.00B	0.27%



Note: Please refer to the metric definitions in the appendix for more clarity

Model Methodology



- **Input**:
  - Customers: Made purchase in the past 3 months.
  - Attributes (from the point of their latest purchase within the 3 Month window) :
    - Past 6 Months data (Sales).
    - 1 year data (Demographic data)
- **Output**: Probability/Segments of the customers to shop within 90 days from the time of their last purchase date;

Example:

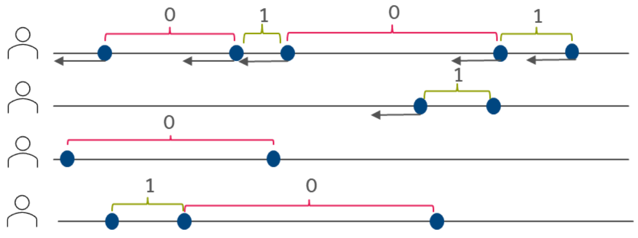
- **Scenario 1**: Customers A & E make an actual purchase within 90 days from their respective last purchase dates and so was predicted by the model with a high probability.
- **Scenario 2**: However, Customer B had a high likelihood to make a purchase but didn't make an actual purchase.

Classification Model Details

Data Information:

Training/ Testing Population:

- Customers who shopped at least once in the past two years.



- Classification Algorithm: XG-Boost Classifier
- Challenges: Class Imbalance
  - Counter:
    - Weight balancing technique
    - Over Sampling/ Under Sampling

Over Fitting:

- Hyper-Parameter Tuning using Bayesian Optimization

Features	Importance
Frequency (Past 6 Months)	32.08%
In-Store RFM Decile	23.02%
# of Fiscal Periods corresponding to purchases	11.87%
# of KH DM Received in the last 90 days	10.21%
Home to Store Distance	7.13%
Sales (Past 6 Months)	6.16%
Marital Status	2.48%
Off-Coupon RFM Decile	1.81%
# of Circular Received in the last 90 days	1.64%
# of Shopped Product groups	1.24%
Presence Of Child	1.10%
ECOM Recency Decile	0.72%
ME redemptions in the last 90 days	0.32%
# of Transactions in previous Quarter	0.22%

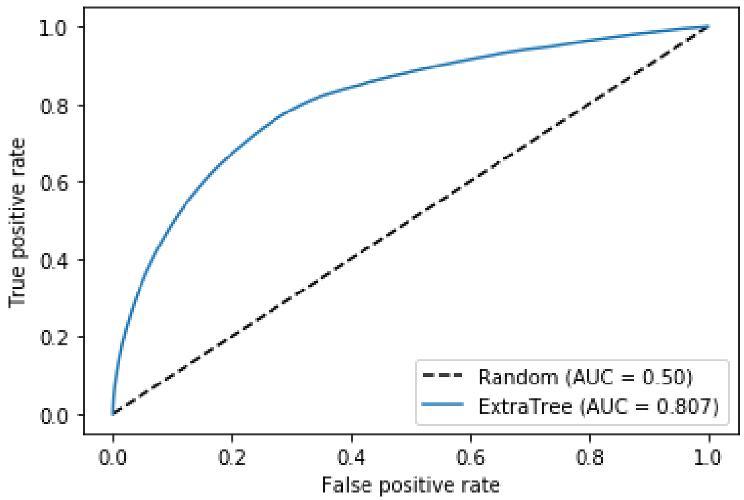
Model Accuracy

Training/ Testing

Training				
	Precision	Recall	F1-score	Support
0	70%	73%	72%	1,766,474
1	78%	75%	77%	2,233,526
Accuracy		75%		4,000,000

Testing				
	Precision	Recall	F1-score	Support
0	70%	73%	72%	1,322,595
1	78%	75%	77%	677,405
Accuracy		75%		2,000,000

ROC Curve of XG-Boost Classifier



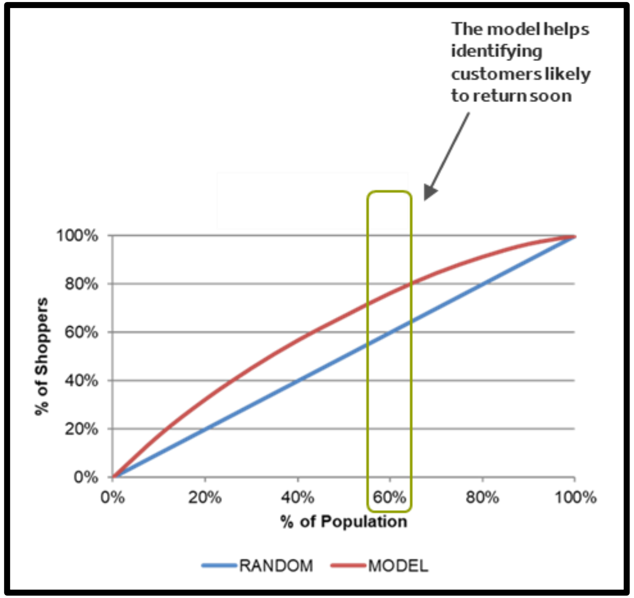
We focus on Recall % as we do not intend to miss on the customers who are most likely to shop

Projected Lift of the Model

- If we Emailed everyone, we would arrive at 100% of possible Shop Rate.

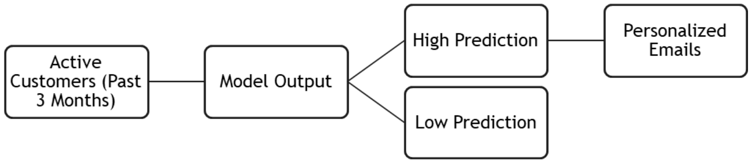
- With the business deciding to mail 60% of the Email Contactable, the graph helps us in investigating what effect this would have.

- **<u>In theory</u>**, Emailing 60% of the population would yield 76% of sales when using the Model vs Random Selection.



Back-testing

Applying Personalized Product Recommender on Active Customers



The output generated from the model could be used in the following way:

- To design a separate marketing strategy to ensure the customer having high shopping probability to make their next purchase.
- Potential for finding customers for up-selling and cross-selling.

Personalized Product Recommender

A next generation ML driven product recommendation solution designed to deliver personalized product offerings based on customer's persona and recent product engagement.
For each customer persona, a set of product recommendations are created based on product-to-product affinity.

Customer's persona is defined by a layer of customer segmentation

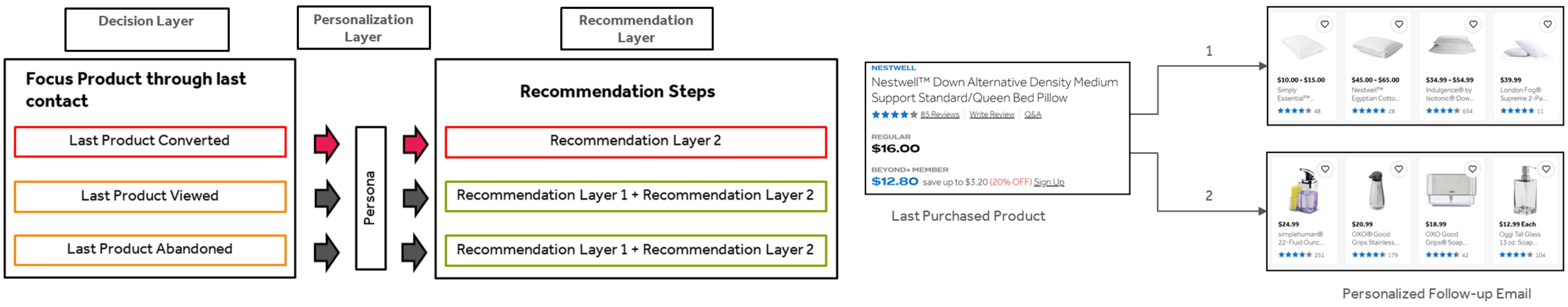
- Using Customer Demographics + Customer Geographic Data + Transactional Data + Life Stage to form a persona, allowing a more granular and precise personalized grouping of customers

Recommendation Flow

Product Recommendation consist of two layers:

- **Recommendation Layer 1 –**This layer intends to recommend products with the highest probability to convert within the same product type for a given customer persona.

- **Recommendation Layer 2 –** This layer intends to recommend other products based on their affinity to the converted product for a given customer persona.



Model Details

Personalization Layer:

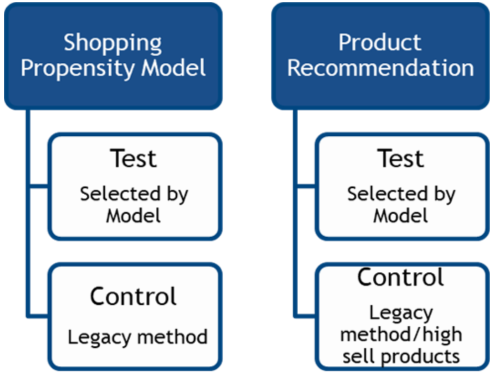
- **Model Training:**
  - Active customers in the past six months.
  - Attributes: Customer Demographics + Geographic Data + Transactional Data + Life Stage during the time of each transaction.
  - Algorithm: **K-Modes Clustering** (Presence of all categorical variables).
  - Clusters/Segments: 6
- **Cluster Prediction**:
  - K-Mode algorithm can back-predict the clusters for unseen customers and customers with changed attributes.

- **Recommendation Layer 2:**
  - **Model Training:**
    - Algorithm: Matrix Factorization.
    - Separate Models for each Personalization Cluster.
    - Data: Sold units of every Product type for each customers belonging to the respective clusters.
  - **Model Output**:
    - Flat file containing ranked product types to every focus product type according to their affinities.

Test Framework - Email

- Shopping Propensity
  - Using ML to identify most likely shoppers
  - Comparing against legacy selection method
  - Success Metric: 2% lift in Shop Rate

- Product Recommendation
  - Predicting most suitable product types to active customers
  - Comparing against high sell product types or legacy method
  - Success Metric: 2% lift in click /open/CTOR



Segment	open rate	CTOR	CTR
A	17.81%	2.73%	0.49%
B	5.06%	15.74%	0.80%
C	1.20%	14.95%	0.18%
C+	1.62%	15.09%	0.24%
D	2.68%	14.53%	0.39%

Segment	Mailed Customers	Customers Shopping	Shopping Rate	Avg. Sale	SPNE	Coupon Expense Rate	MPNE	OPEN RATE	CLICK RATE	CLICK TO OPEN RATIO
Test	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Control	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
	Net Diff.		2.06%	1.24%	2.52%	0.06%	2.04%	1.24%	0.97%	2.29%

*<u>Representative Sampling Objective:</u> Test & Control must be comparable to make sure results are valid
*<u>Need for test:</u> To know how much we put at risk or stand to gain at scale

More in Pipeline

- Implementing Personalized Product Recommendation through Direct Mails and other marketing channels.
- Creating a Multi-Channel Propensity Model to stream-line Marketing efforts through appropriate channels.
- Testing Strategy for customers having lower probabilities with product recommendations and higher discounts.
- Creating an ML Solution for Lapsed and At-Risk Customers.

Thank You!

Sample Data

Last Product Converted → Recommendation Layer 2

Last Purchased - Layer 2		
dw-bq-data-p00:ANALYTICAL.customer_last_purchase_recommendations		
Description: Recommends top 3 product types based on affinities, to each product type top 5 SKUs are recommended based on the probabilities for the recommended product type, persona		
Field name	**Description**	**Type**
customer_id	Customer ID	INTEGER
last_purchase_day	Last Purchase Date	DATE
item_sku_num	SKU Number of the focus Product	INTEGER
item_product_type_id	Product Type ID of the Focus Product	INTEGER
persona	Customer Segment (Persona)	STRING
recomm_pdm_prod_type_id	Product Type ID of the Recommended Product	INTEGER
product_type_rank	Affinity Rank for Product Type	INTEGER
layer2_rec_sku_num	SKU Number of the Recommended SKU	INTEGER
sku_rank	Probability Rank of the recommended SKU for Recommended Product Type for specific Persona	INTEGER
sku_rank2	Probability Rank over Focus Product for the recommended SKU for specific Persona and Sub-Persona	INTEGER

Sample Data

Focus Product SKU	Focus Product Type	Recomm Product Type	product type rank	LAYER2 REC SKU NUM	LAYER2 REC SKU	SKU Rank	SKU Rank3	SKU Rank2
CUIS CHF CLSC PRO SS 11PC SET	Cookware Set	Pot Holder	1	43875726	AC PEWTER PH	1	30	10
				43875795	AC CORNFLOWER PH	2	38	12
				43875733	AC BLACK PH	4	40	13
				43875740	AC CHILI PH	3	41	14
				69465529	KSMART GREY PH	5	44	15
		Bath Rug	2	60544339	SE MEM FOAM CHARCOAL 21X34 VDS	1	14	6
				60544407	SE MEM FOAM MEDIUM BLUE 17X24	2	21	7
				60544414	SE MEM FOAM MEDIUM BLUE 21X34	4	23	8
				60544445	SE MEM FOAM CHROME 17X24	5	25	9
				60544285	SMART DRY DEEP LINEN 17X24	3	34	11
		Food Container	3	63762495	FOOD STRG OXO POP 2 WT 1.7Q RC	1	4	1
				63761528	FOOD STRG OXO POP 2 WT 4.4Q SQ	2	5	2
				63761535	FOOD STRG OXO POP 2 WT 2.8Q SQ	3	6	3
				63761504	FOOD STRG OXO POP 2 WT 1.1Q SH	4	10	4
				63762488	FOOD STRG OXO POP 2 WT 2.7Q RC	5	13	5
INTRIGUE 18X30 MAT AQUA	Kitchen Mat	Door Mat	1	60182272	SUPER SPONGEENTRYCHAR21X34	2	15	3
				60182289	SUPER SPONGE ENTRY BRWN 21X34	1	16	4
				40732176	DOORMT TRAPR 17X29	3	19	6
				44937140	ARGYLE CAMEL 20"X30"VDC	4	24	8
				43312108	DOORMT CLEAN TURF	5	27	9
		Drapery Rod (Single)	2	61519046	CAFÉ ROD BALL NICKEL 28-48	1	29	10
				61519053	CAFÉ ROD BALL NICKEL 48-84	4	36	12
				17157558	PREM BN 48-88	2	37	13
				16471690	TENSION OV 22-36 WHT	3	42	14
				17157507	PREM BN 28-48	5	43	15
		Frying Pan	3	69547083	ALLCLAD 2PK NS FRY 10&12	1	2	1
				68054861	TFAL PURE CK FRY PAN 3P BLK	2	12	2
				67417131	BLUEDIAMOND FRY PAN 2PC	3	17	5
				69509165	CALPH PREM FRY 2PK	4	20	7
				62717915	BIALETTI TITAN FRY PAN 10	5	35	11
INVIT SQR GRDL 10.5 PP GRY	Frying Pan	Window Panel (single)	1	69497045	VOILE SHR WHITE 84	1	18	7
				47021457	QUINN IVRY 84	2	22	8
				47021495	QUINN WHT 84	5	26	9
				47021464	QUINN LIN 84	3	31	11
				17967363	CRUSHDVLE PL WHT84	4	33	13
		MICRO BOWL S/4 SALT	2	65114834	MICROWAVE MAT 12" PREP SOL	1	3	2
				11970680	MICRO PLT S/4 SALT	2	28	10
				44135454	MICRO BOWL S/4 SALT	3	32	12
				69644010	MICRO PLATE S/4 SIMPESSEN	4	39	14
				13461406	BACON CKR MICRO SALT	5	45	15
		Scent Packet	3	14429689	SCENT PKT WHT CTN 3	1	1	1
				12160585	SCENT PKT LAVNDR 3	2	7	3
				40155937	SCENT PKT SMMR RMC 3	5	8	4
				18082705	SCENT PKT WTRMRK 3	4	9	5
				12042620	SCENT PKT SEASPRY3	3	11	6