Using **Data Science** to make Your Online Retail a **Success**

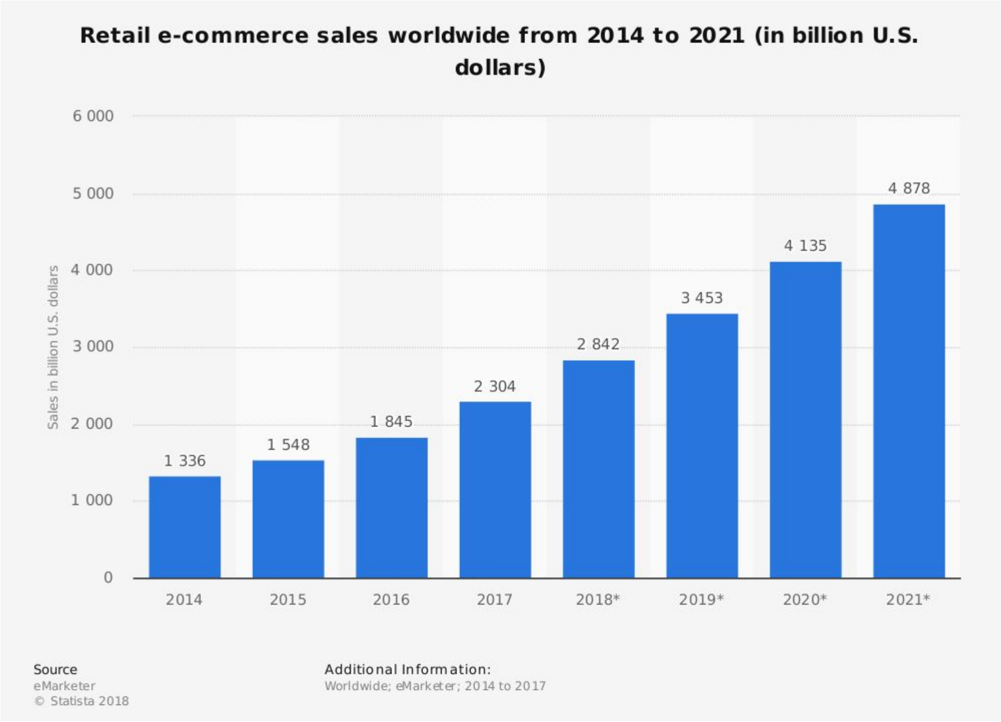
BAX 453 Application Domains

Final Project Report

Team Members: Leo, Maggie, Nicholas, Siyu, Zimei

May 30th, 2018

Congratulations! You’ve chosen a boosting industry!

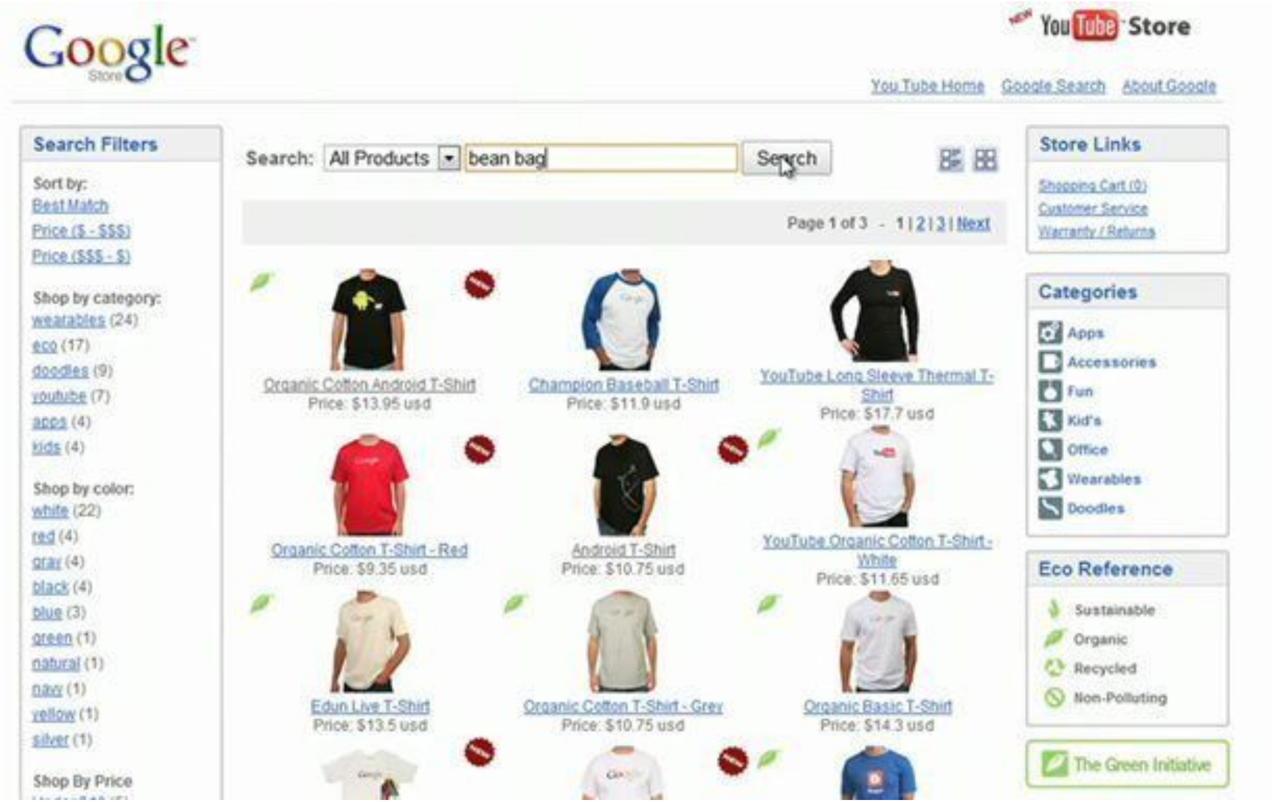


But, that also means competition...

What do you do?

Offline Shopping Experience

Online Shopping Experience

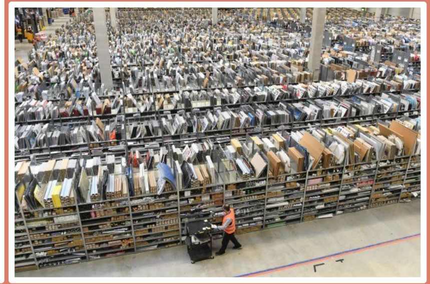


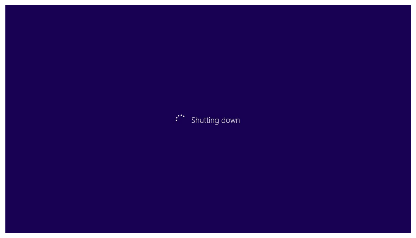

ONLINE SHOPPING IS TOO HARD
THERE'S TOO MANY CHOICES
quickmeme.com

Shutting down

Pleasant In-Store Experience

Frustrating Online Shopping





Solution!!! **Recommendation Engine**

A Use Case:

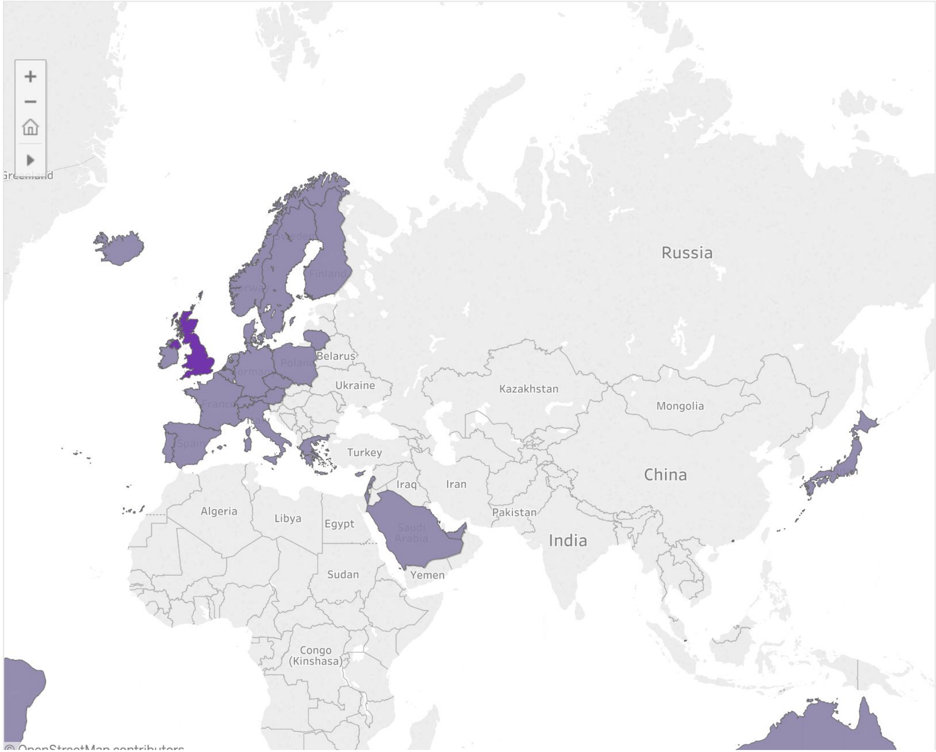
Implement Recommendation Engine for UK Online Retailer

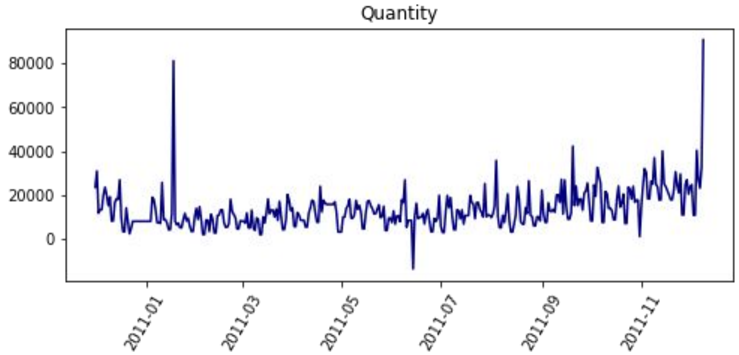
Types of Data Required

We will use transnational data set which contains all the transactions occurring between 01/12/2010 and 09/12/2011 for a UK-based and registered non-store online retail.The company mainly sells unique all-occasion gifts.

- InvoiceNo
- Description
- Quantity
- InvoiceDate
- UnitPrice
- CustomerID
- Country

Data Exploration





Algorithm

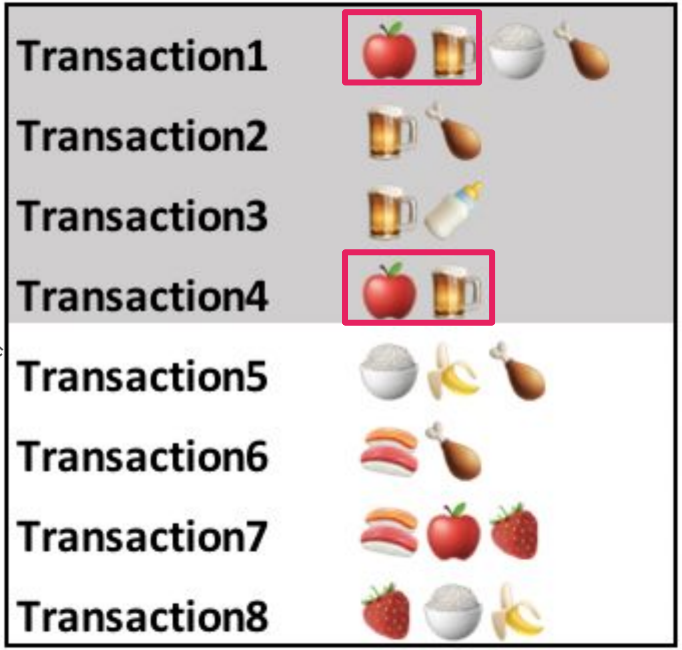
Support of {🍺, 🍎} = 25%

Apriori algorithm - *"A priori"*

Association rules mining

Measure 1. **Support** $support(A \Rightarrow B) = P(A \cup B)$

How popular an item is, measured the proportion of transactions where an item is purchased.



Support of 🍺 = 50%

Algorithm

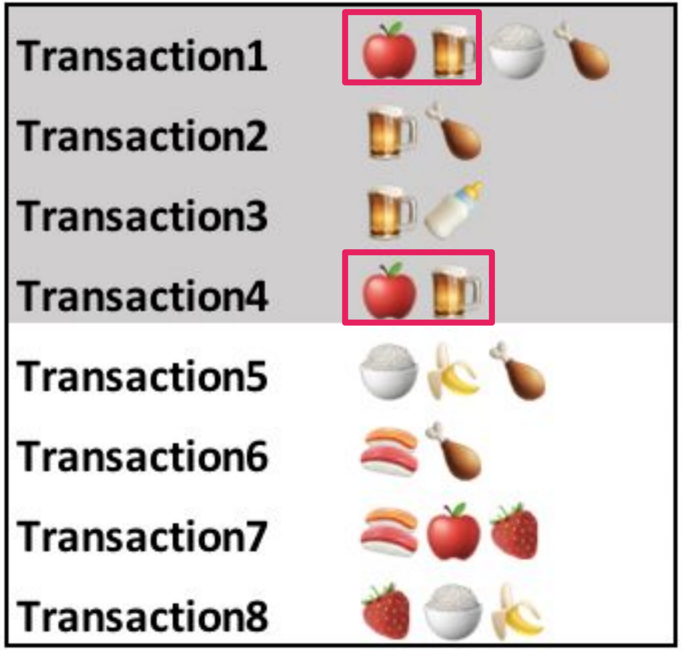
Apriori algorithm - Association rules analysis

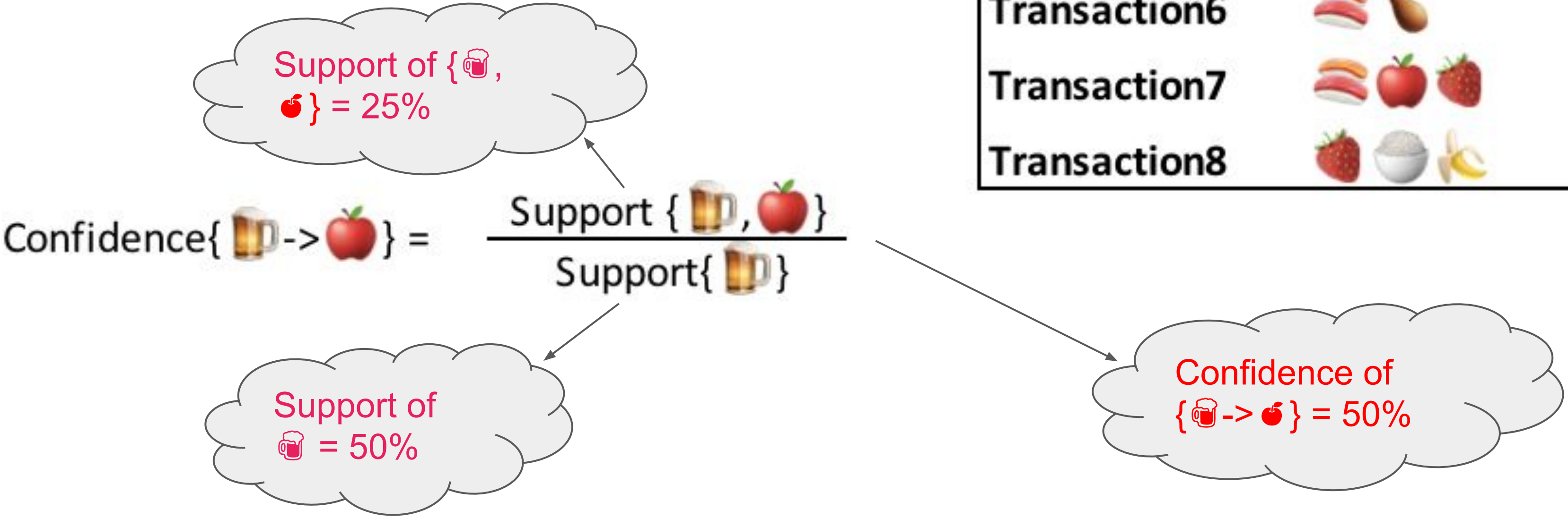
$$confidence(A \Rightarrow B) = P(B \mid A) = \frac{support(A \cup B)}{support(A)} = \frac{support\_count(A \cup B)}{support\_count(A)}$$

Measure 2. **Confidence**

How likely item Y is purchased when item X is purchased, denoted as {X ->Y}.

Proportion of transaction with item Y, in which item X also appears.



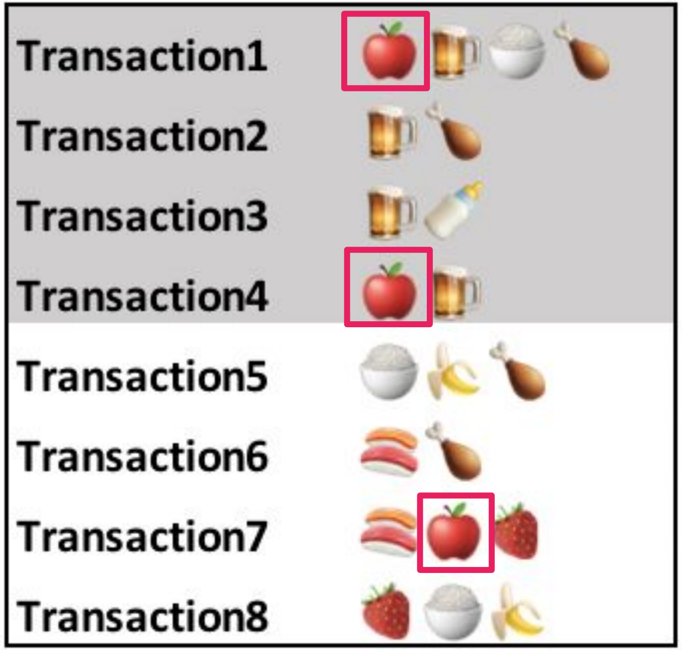


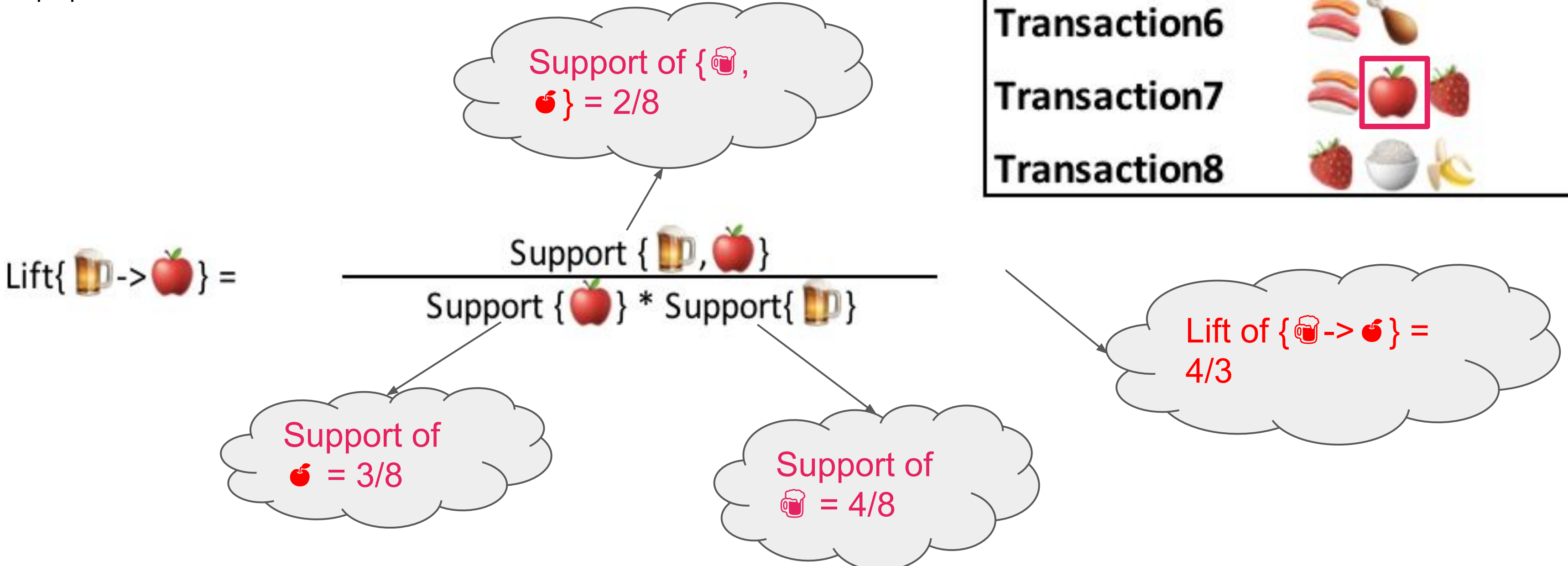
Algorithm

Apriori algorithm - Association rules analysis

Measure 3. **Lift**

How likely item Y is purchased when item X is purchased, while controlling how popular item Y is.





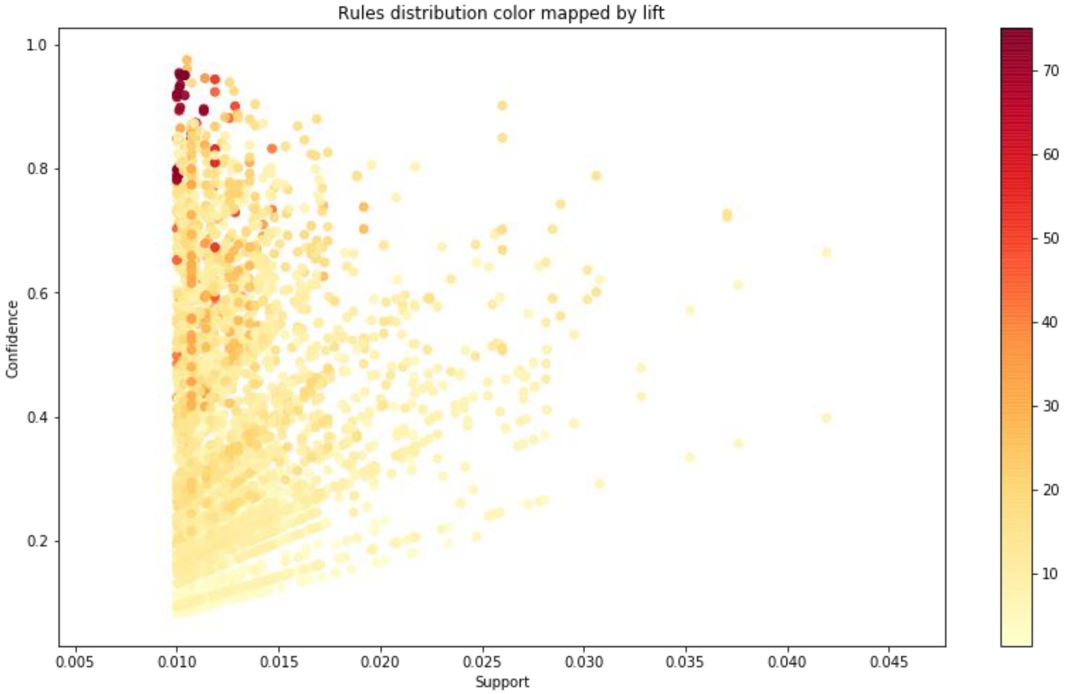
Results

X (antecedents)	Y (consequents)	support	confidence	lift
HERB MARKER MINT, HERB MARKER THYME	HERB MARKER ROSEMARY	0.01	0.95	75
KNITTED UNION FLAG HOT WATER BOTTLE	RED WOOLLY HOTTIE WHITE HEART	0.06	0.65	5.58
SET OF 3 CAKE TINS PANTRY DESIGN	JAM MAKING SET WITH JARS	0.02	0.36	5.02

- Support: This combination of the two items takes 6% part of all the itemsets
- Confidence: The white heart bottle is 65% likely to be purchased when a union flag one is purchased
- Lift: Lift > 1. If we control how popular white heart one is in the market, we can still say it’s very likely to be bought if union flag one is bought.

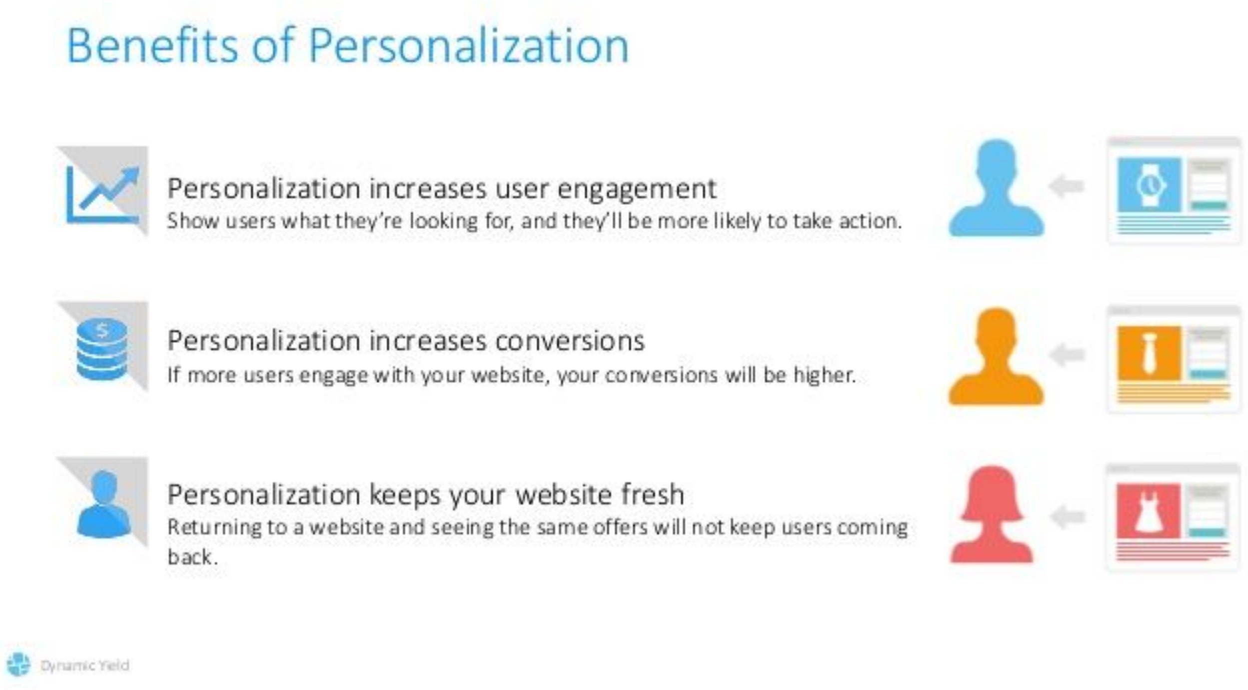
Metrics to Measure Model Performance

- Support, Confidence, and Lift of the Association Rule
- Gross Merchandise Volume
- Revenue Gain
- Number of Active User
- User Purchase Frequency
- Average Basket Size (in quantity and in $)
- Average spending in $ per user (to find the most valuable customers)



Application

Recommendation engines enable personalization of online retail



Conduct Cross-Sell and Up-Sell Campaigns to Increase Sales

Cross-sell and up-sell campaigns show the products purchased together, so customers who purchase the Samsung monitor can be persuaded to pick up high quality HDMI cable.

Frequently bought together

Total price: $239.98

Add both to Cart

Add both to List

- **This item:** Samsung UE510 LED DISPLAY Monitor, Black, 28" 4K (Certified Refurbished) $229.99
- HDMI Cable 6ft - HDMI 2.0 (4K @ 60Hz) Ready - 28AWG Braided Cord - High Speed 18Gbps - Gold Plated... $9.99

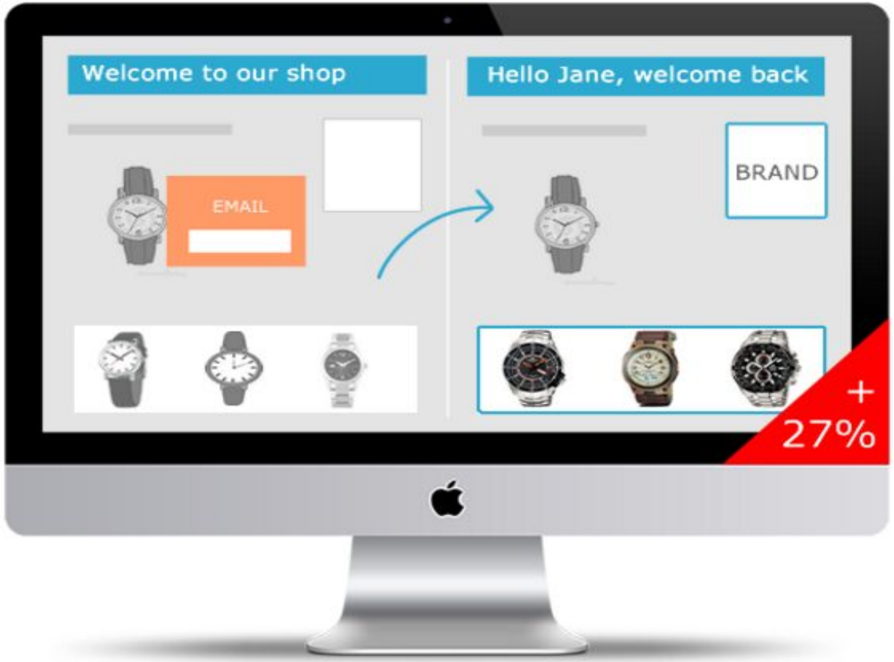
Design Effective User Interface and Product Combo Offers

Product combos are based on product affinities, developing combo offers and design effective user friendly interface by focusing on products that sells together.

Quantity	Buy 1 Get 1	Combo	Free Gift

Build Shoppers Profile to Optimize Customer Experience

Shoppers profile in analyzing market basket with the aid of data mining over time can help retailers to get a glimpse of who their shoppers really are, gaining insight to shoppers' spending range, buying habits, likes and dislikes, and purchase preferences, and optimizing customer experience based on the information retrieved.



Challenges of Recommendation Engine in E-Commerce

Data and Product Dependency

- Recommendation engines' knowledge of a person is based on user activity within an interface and purchase records, the efficacy of the algorithm itself, and the wisdom of the crowds.

- Although some users' behavior can be modeled, other users do not exhibit typical behavior. These users can skew the results of a recommender

Lack of Innovation

- In trying to select a winner product, recommendation engines tend to reproduce stereotypes and reinforce existing practices.

- Recommendation algorithms usually don't support the “Long Tail” enough and just recommend obvious items.

- While algorithms can learn from weighted variables in a product database, the listing itself often has a human touch, including the product description, image, and overall web design. A recommendation engine may thus favor one product over another simply due to better imagery and not product quality.

Changing User Preferences

- User preferences might change rapidly. while today some users have a particular intention when browsing on an online retailing website, tomorrow they might have another intention.

- Especially, in some item categories, for example, fashion clothing, user preference can be totally different after a short period of time, due of the change of fashion trends. Recommendation engines are usually behind the changing trends and take quite some time to update and pick up the changing user preferences.

Filtering Useful Rules

- A problem with recommendation engines based on market basket analysis is that sometimes too many rules are generated and it becomes important to filter these rules to select the strongest or the most relevant.

- The bottom line is that for performing an efficient market basket analysis, simply applying the algorithm on the available data may result in a profusion of association rules. If we are not careful about how we apply these rules we may lose some valuable information.

Reference

IndraStra Global. (2016, March 08). IT | Benefits and Challenges of Data Mining in E-Commerce. Retrieved from https://www.indrastra.com/

Five Problems of Recommender Systems. (2009, January 29). Retrieved from https://readwrite.com/2009/01/28/5_problems_of_recommender_systems/

Deshpande, B. (n.d.). Challenges in filtering useful rules from a market basket analysis. Retrieved from http://www.simafore.com/blog/bid/113493/Challenges-in-filtering-useful-rules-from-a-market-basket-analysis

Market Basket Analysis and Recommendation Engines. Retrieved from https://www.knime.com/knime-applications/market-basket-analysis-and-recommendation-engines

Thank you!


Woohoo!