RightFit

Using Data Science to Reduce Returns Due to Incorrect Product Size

The Fifth Elephant, 2016

Ashish Kulkarni
Principal Data Scientist, Jabong Labs
https://in.linkedin.com/in/ashishakulkarni

Outline

- Motivation and Introduction
- What’s out there
- RightFit - This is what we did
- Evaluation

Outline

- **Motivation and Introduction**
- What's out there
- RightFit - This is what we did
- Evaluation

Why do we care?

Lot of fashion products (apparel, shoes, etc.) get either exchanged or returned

Most of these are attributed to incorrect size (fitment) issue

This adds to the cost of reverse logistics, customer dissatisfaction, marred experience and their potential churn!

Can data science help?

Let's start at the beginning…

made-to-measure

American civil war
(ca. 1800)

ready-made


source: http://museum.nist.gov/

A union soldier wearing a
standard-size army uniform

"History is filled with brilliant people who wanted to fix things and just made them worse." - Chuck Palahniuk

From couture to the "American look"

	event	sizing standard
	made-to-measure ↑	
1930		
	ready-made ↓	
1940	first sizing survey Ruth O'Brien & William Shelton	- bust - height (regular/long/short) - girth (regular/stout/slim)
1958	NBS (now NIST) publishes commercial standard (CS)	- size (8-38) - height (T, R, S) - girth (+ / -)
1983	Govt. withdraws CS	
	Voluntary standards ↓	ASTM standards
ca. 2000	$[TC]^2$ SizeUSA scanning survey	

hour-glass sample | pear-shaped population

Manufacturers continue to resist ASTM standards

Back to chaos

"The Laura Ashley woman is different from the Liz Claiborne woman, who is different from the woman whom Calvin Klein envisions." - *Times*

Enter "fit" models!

source http://www.popsugar.com

What makes fit so difficult?

- **lack of standardization in cross-brand sizing:** brands design clothes differently using proprietary style specifications
- **vanity sizing (size inflation):** brands size the product lower to appeal to customers' ego

No two sizes feel the same across brands!

Here’s an example…

Phosphorus

SIZE GUIDE

SIZE	Garment Chest (in.)	Garment Length(in.)
39	41	29
40	42	29
42	44	30
44	45	31

Raymond

SIZE GUIDE

SIZE	Garment Chest(In)	Garment Length(In)
39	44	31
40	45	31
42	47	32
44	49	32

Outline

- Motivation and Introduction
- **What's out there**
- RightFit - This is what we did
- Evaluation

Fitting solutions (offline)

Mannequins

Real models

3D solutions

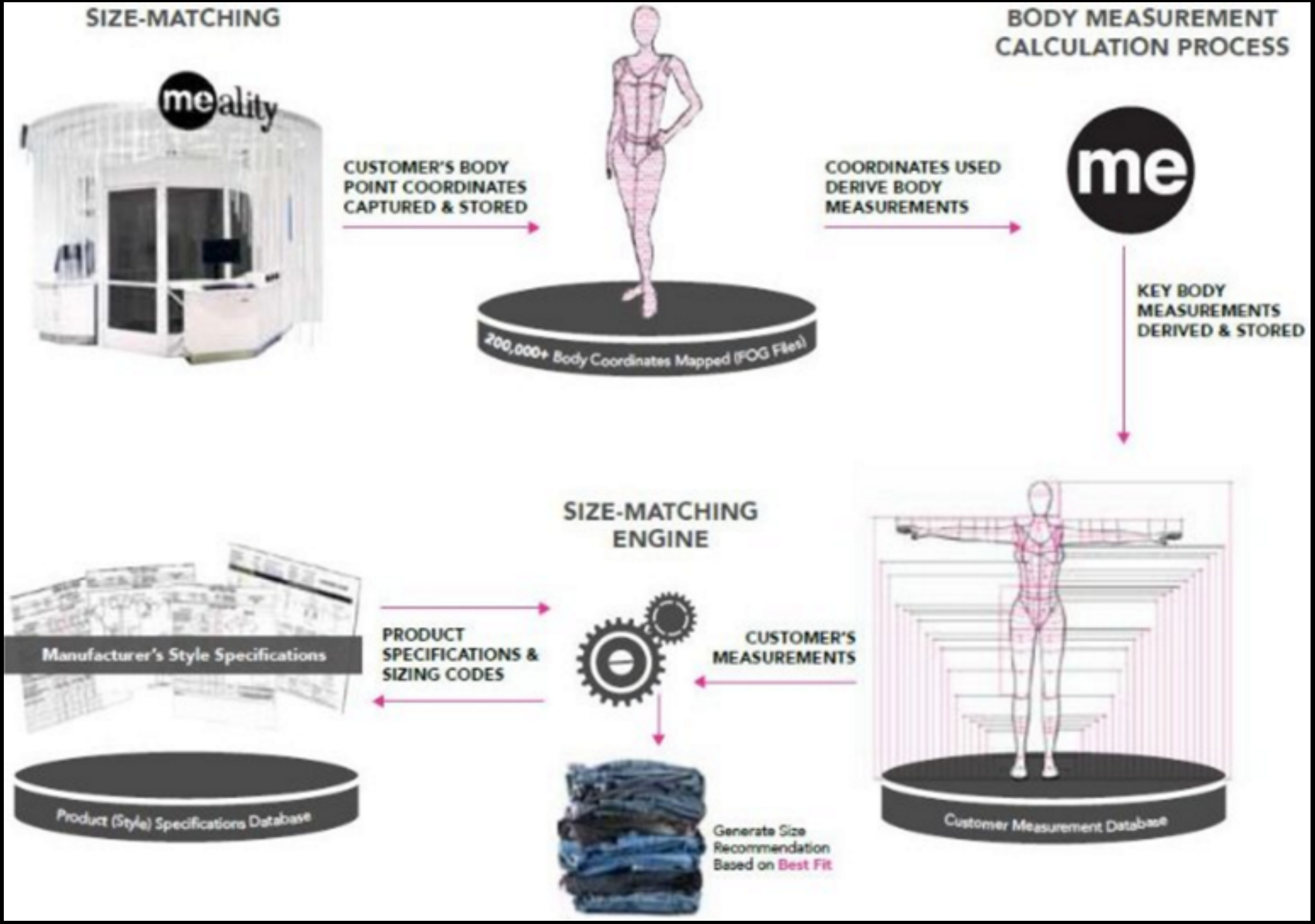
Augmented reality

Virtual fitting rooms

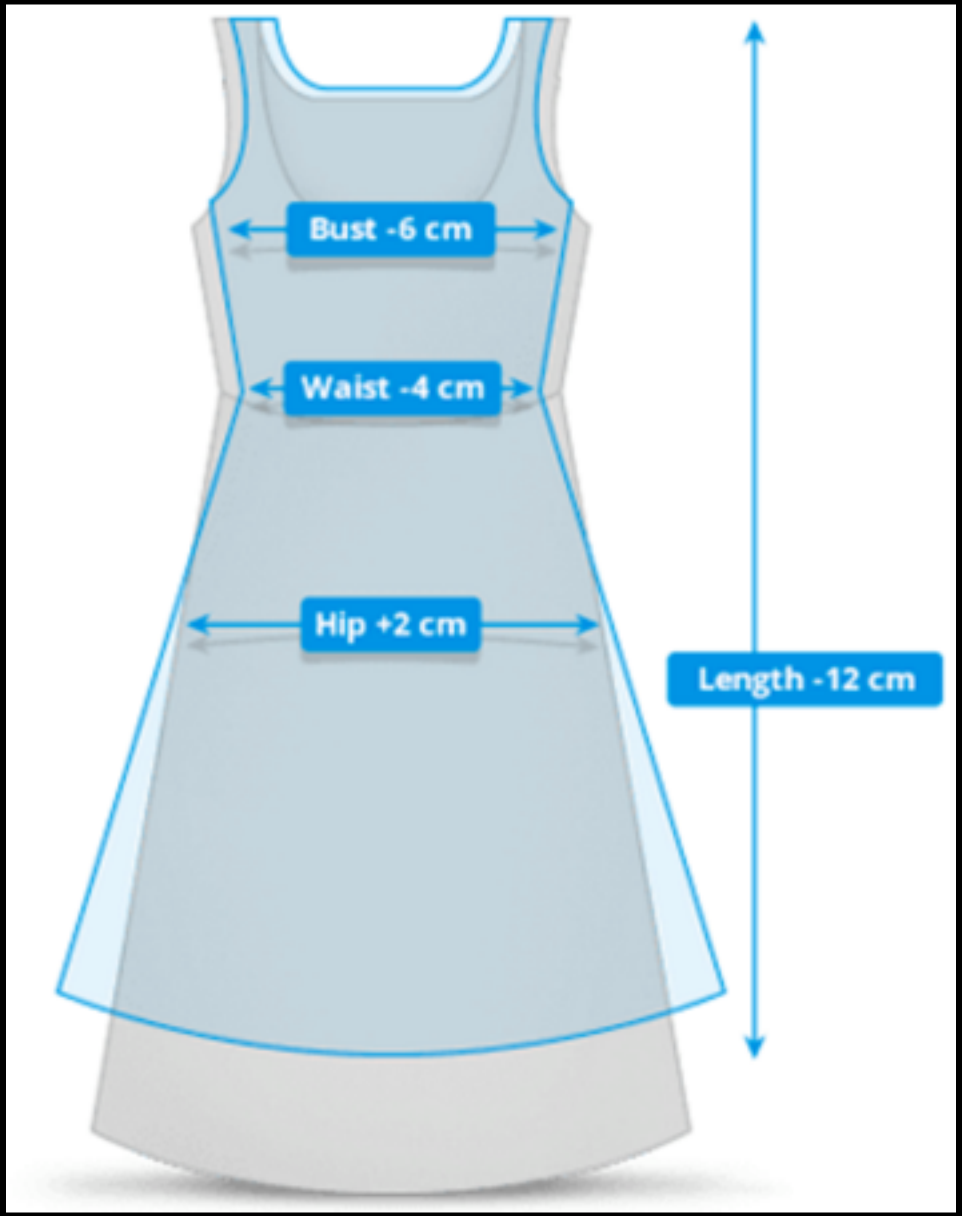
Body scanners

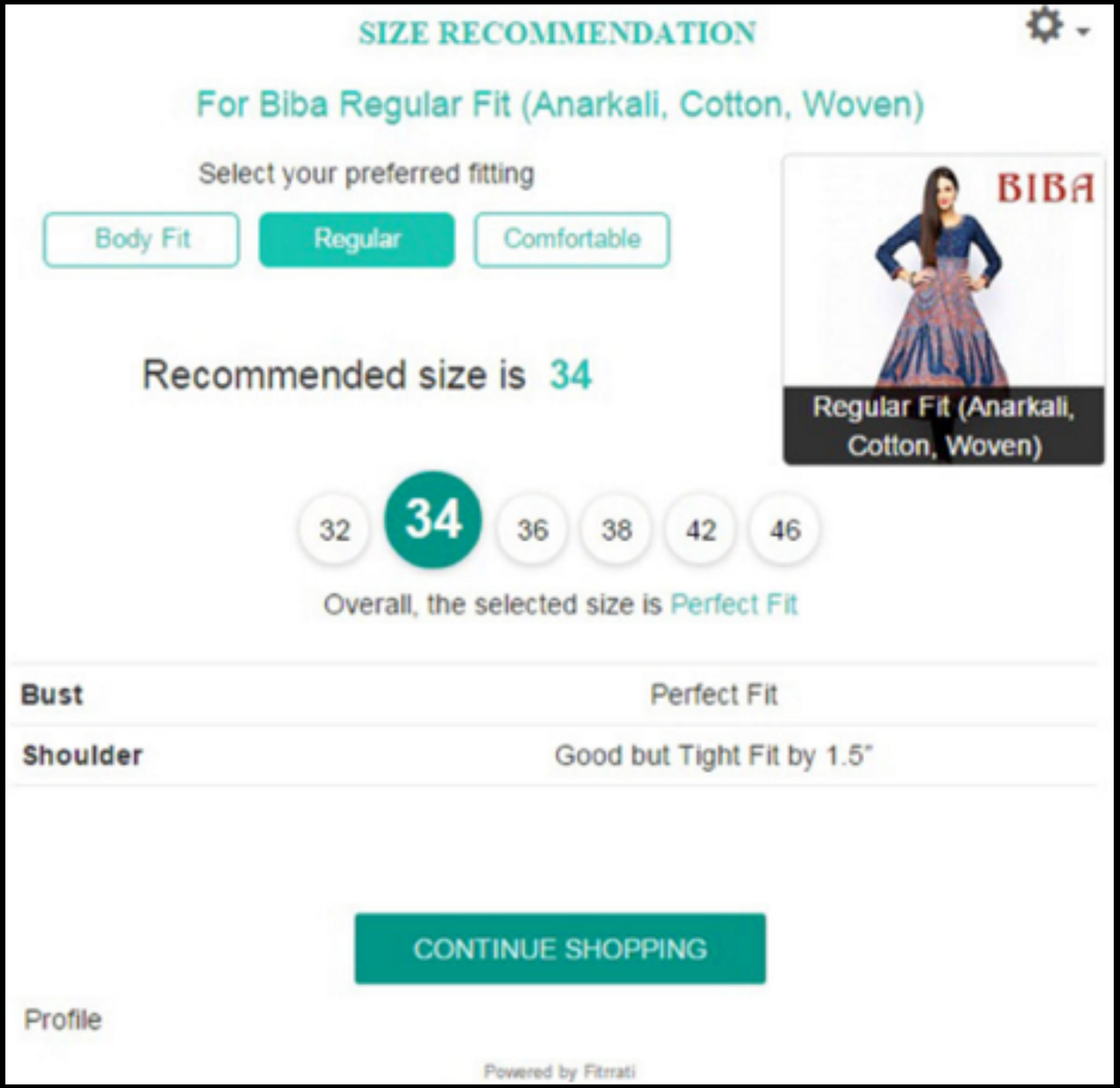
User body measurements → Product size

An example - Me-ality



Fitting solutions (online)

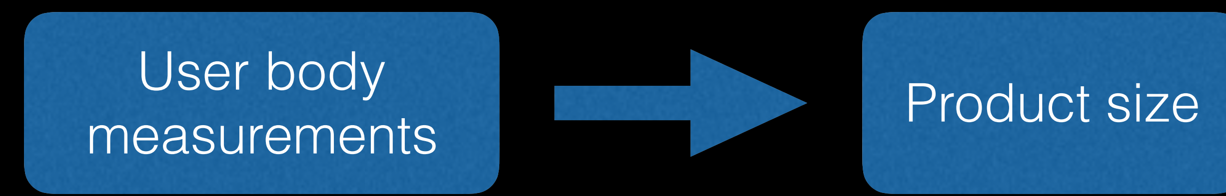




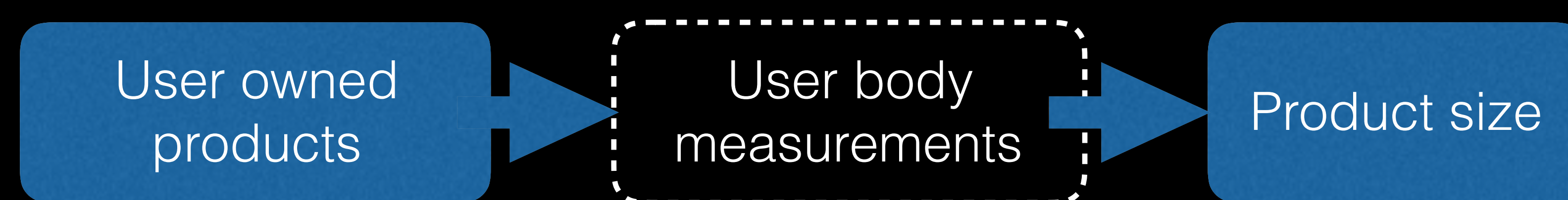
User body measurements → Product size

How do we obtain user body measurements?

- Explicit: Ask users for their body measurements



  - Might be more accurate
  - not scalable
  - bad user experience
- Implicit: Can we infer?



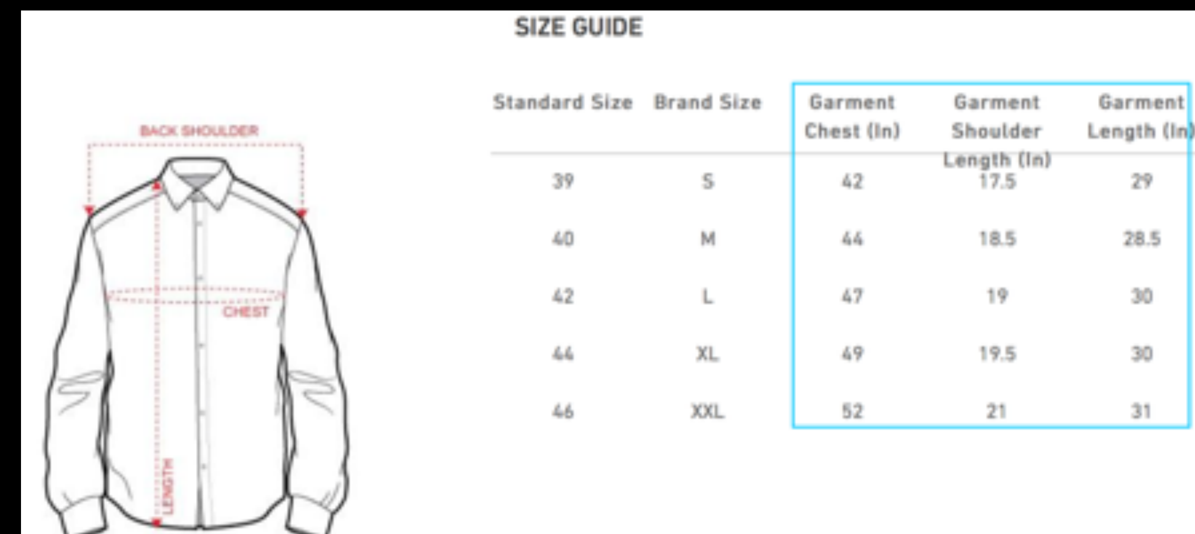
  - what’s in your wardrobe? (first-timers)
  - past buy / return transactions (returning users)

Outline

- Motivation and Introduction
- What’s out there
- **RightFit - This is what we did**
- Evaluation

Our approach

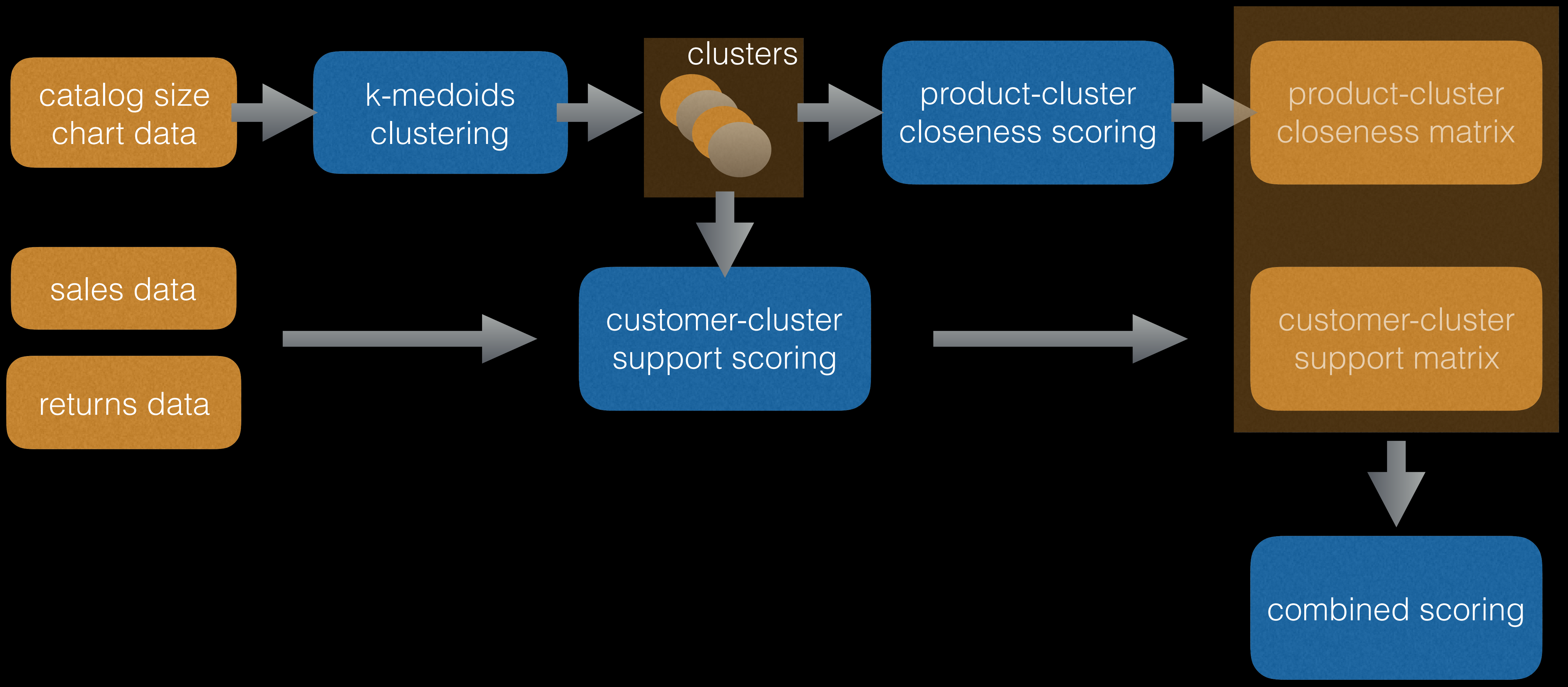
- Sizing standard-based data is noisy; use the **actual size** instead
- A successful sale not resulting into a return / exchange is an indication of right fit
- A return / exchange marked with “size issue” is an indication of incorrect fit



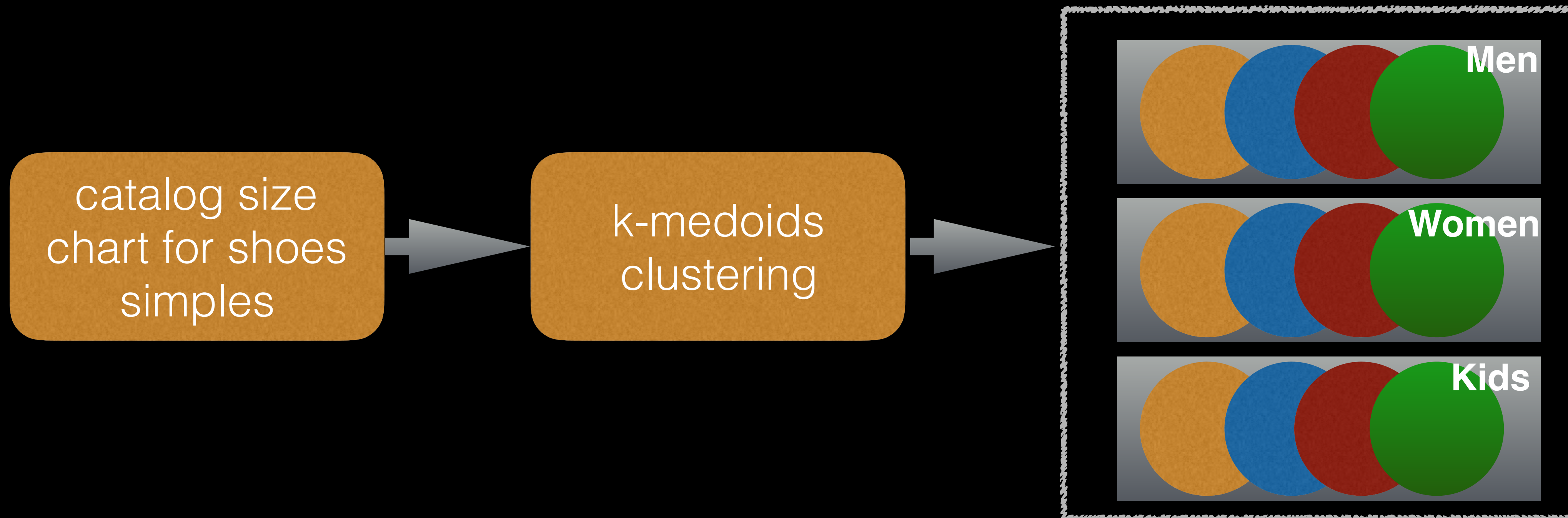
Standard Size	Brand Size	Garment Chest (In)	Garment Shoulder Length (In)	Garment Length (In)
39	S	42	17.5	29
40	M	44	18.5	28.5
42	L	47	19	30
44	XL	49	19.5	30
46	XXL	52	21	31

UK Size	Euro Size	In Cm
6	40	26.2
7	41	27
8	42	27.9
9	43	28.7
10	44	29.6
11	45	30.4

Overall design



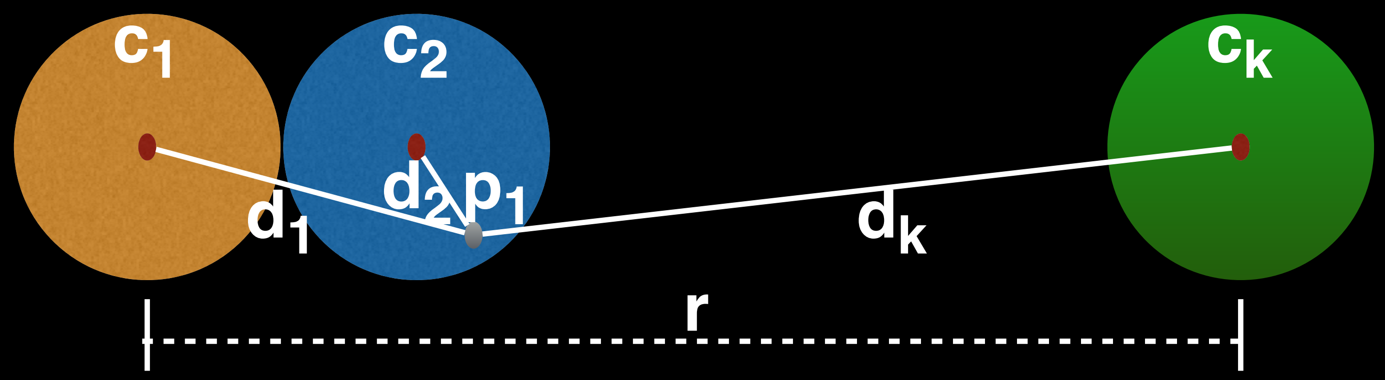
Catalog clustering



clusters of shoes simples

- **k-medoids clustering**
  - used to partition the catalog data into groups of similar sized shoe simples
  - uses “size in cm” feature and squared Euclidean distance as the distance measure
  - minimizes sum of pairwise dissimilarities
  - uses an implementation of **clara** - a variation of PAM (Partitioning Around Medoids) algorithm for large scale clustering
  - The choice of ‘k’ is a function of the range of sizes of shoe simples; currently set to 11 for men, 10 for women and 18 for kids.

Product-cluster closeness scoring



product-cluster closeness score between shoe simple p_1 and cluster c_1 is computed as: $\mathbf{s_{11} = 1 - d_1 / r}$, where, d_1 is the Euclidean distance of the simple from the medoid of cluster c_1 and r is the maximum distance between the clusters and acts as a normalizer.

- The above scoring is used to compute the **product-cluster closeness matrix** comprising closeness scores for all shoes simples from the clusters;
- There is one such matrix each for men, women and kids.

	c_1	…	c_k
p_1	s_{11}		s_{1k}
…			
…			
p_n			s_{nk}

Customer-cluster support scoring

customer	simple	cluster	sales_count	score
x_1	p_1	c_2	4	0.67
x_1	p_i	c_j	1	0.33
…	…	…	…	…
x_m				

	c_1		c_k
x_1			
x_m			

- From the item sales data, we filter out the items returned or exchanged, to retain only the successful item sales;
- This data is augmented with the item cluster membership (Refer to ‘cluster’ column) by combining it with the output of our clustering;
- We then group this data by *customer* and *cluster* to obtain *sales_count,* the number of times a customer bought a shoe simple from that cluster;
- Finally, the **customer-cluster support score** is computed as:

(sales_count + 2) / (sum(sales_count) + 4), where, sum(sales_count) is the total number of sales for the customer and the score is adjusted as per Wald’s method to account for sparseness.

Support-closeness combined scoring

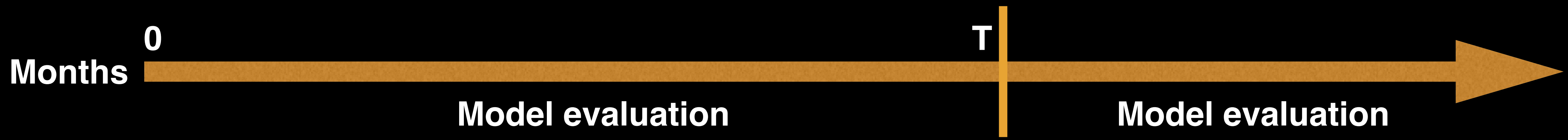
- For a given SKU, get the closeness vectors for each of its simples to create a closeness matrix *M*;
- For a given customer, get the support vector *V* from the customer-cluster support matrix;
- The element-wise multiplication *M * V,* gives the combined scores;
- We apply row-wise max on *M * V* and order the simples by decreasing scores.

- In the absence of information on which "profile" (cluster) the customer is buying for, we return the simple with the maximum score;
- If we have the knowledge of his/her profile, we return the simple with the max score in the column corresponding to that profile (cluster) in the final matrix.

Outline

- Motivation and Introduction
- What's out there
- RightFit - This is what we did
- **Evaluation**

Evaluation



- Successful sales data (after excluding the returns) between 0 to T months was used for model building;
- Successful sales and returns data for the next two months (T+1 to T +2) was used for model evaluation;
- The model was evaluated on two metrics - how precisely does it recommend a right size (leading to a successful sale) and its accuracy in identifying a potential return;
- We performed three runs with different values of T. On an average, the model was 88% precise in recommending a right size and was about 60% accurate in identifying a potential return.


Thank you