FIT PREDICTIONS

minimize returns | maximize revenue

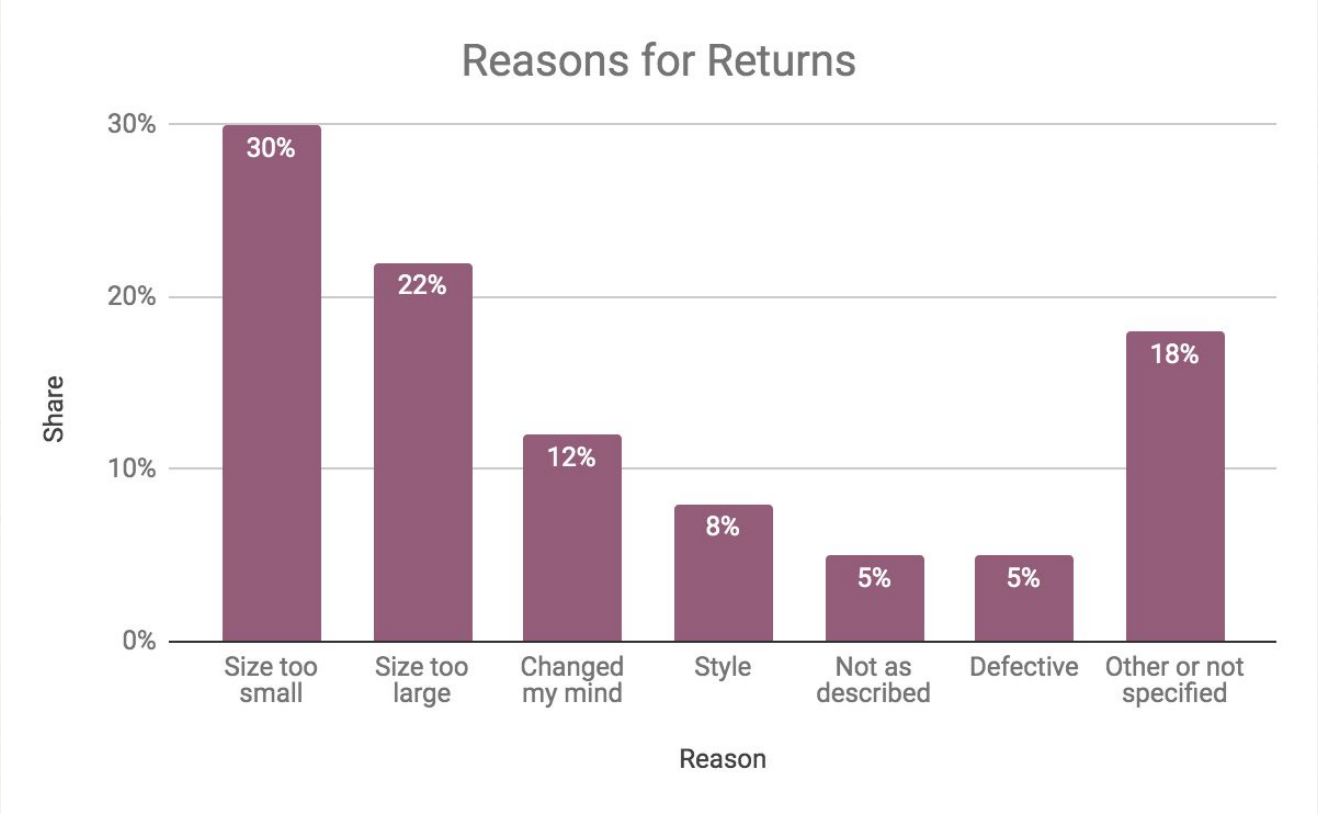
DSI Capstone Project

Arielle Miro

Consumer preference-based return reasons (e.g., size, fit, style, etc.) tend to drive around 72% of all returns in fashion product categories..

Source: https://www.shopify.com/enterprise/ecommerce-returns

CURRENT RETURNS STATS



Incorrect sizing accounts for over 50% of returns in the e-commerce retail space.

Source: https://www.shopify.com/enterprise/ecommerce-returns

1. PREDICT “FIT”

Battle the plague of return rates through smart sizing predictions

DATA FRAMEWORK

Rent the Runway measurements, user info, and review information.

SIZE & CATEGORY

- User
- Item
- Category
- Occasion
- Size
- Body Type
- Weight
- Height
- Age

REVIEWS

- Review Summary
- Full Review

Link to EDA:

https://public.tableau.com/profile/arielle7797#!/vizhome/Capstone_Rent_the_Runway/Understandingthedata

https://public.tableau.com/profile/arielle7797#!/vizhome/Capstone_EDA_2/AdditionalEDA

PREDICTIVE MODELS

Multiclass Classification

SIZE & CATEGORY

- All Numeric Columns
- Random Forests
- Best Score: 98.16%

REVIEWS

- NLP
- CountVectorizer
- Random Forests
- Best Score: 98.54%

CHALLENGES: Unbalanced Classes → Smote → Gradient Boosting

MODEL IMPROVEMENTS

2. RECOMMEND

Increase revenue & enhance user experience with smart product recommendations

COLLABORATIVE RECOMMENDERS

USER BASED

Recommends based on similar users. Assumes that similar users have similar tastes.

ITEM BASED

Recommends based on similarities between items.

CHALLENGES: User profiles in this case are not based on preference. The data does not have item details.

CONTENT RECOMMENDER

Enter the index

```
1 recommendations(29207)
```

```
[1867, 26019, 8835, 4167, 2198, 11384, 16851, 26539, 7122, 5245]
```

Search for the corresponding indices

```
df[df.index == 1867]
```

```
df[df.index == 26019]
```

View recommendations:

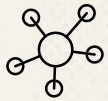
	age	body_type	bust_size	category	fit	height	item_id	rating	rented_for	review_date	review_summary	review_text	size	user_id	weight
1867	44	hourglass	34c	shirtdress	1	5. 4	1840637	6	party	2016-08-05	Very cute and preppy	This dress is so cute! I've been wanting to re...	4	46348	108

	age	body_type	bust_size	category	fit	height	item_id	rating	rented_for	review_date	review_summary	review_text	size	user_id	weight
26019	39	petite	32d	gown	1	5. 2	141688	8	formal affair	2014-02-09	Beautiful dress and even better in person!	This was my first experience with Rent the Run...	12	401375	123

NEXT STEPS

Build “Master” Model

Leverage AWS to model the combined data for sizing and reviews. Compare this score to the separate models.



Expand on Content Based Recommender

Given the shortcomings of the collaborative recommenders, incorporate NLP on review data for item sentiment.

ADDITIONAL IMPROVEMENTS

Purchase Amount

To perform predictive analytics on revenue projections. Also, calculating revenue loss due to returns.

User Location

Additional profile detail to analyze the consumer demographic beyond age and size.

Item Description

Additional information about the item will allow for more granular predictions. An image of the item may allow for further image recognition comparisons.

THANKS!

Any questions?