

## Group 1 | Final Project | Ecommerce Consumer Behavior

### Analysis Summary | Segment 2

#### 1. Products Recommendations | Apriori Algorithm

- a. **Question:** Can we predict which products a customer will most likely purchase together within various product segments?
- b. **Goal:** Increase conversion rates and sales (cross sell) based on recommendations of products bought together
- c. **Preliminary Insights:**
  - Higher confidence in products bought together in the same product category instead of various categories
  - Videos and Music categories with list of recommended products with higher confidence (items with 70% confidence), other categories have lower confidence levels:
    - Video games: 55% or lower
    - Watches: 44% or lower
    - Office products: 43% or lower
    - Apparel: 35% or lower
    - Personal care appliances: 16% or lower
    - Furniture: 13% or lower
  - Preliminary results with higher confidence level appears to be plausible:
    - For video: VHS for Bambi, Peter Pan, Aladdin, Beauty and the Beast, Lady And the Tramp, Little Mermaid and Cinderella.
    - For music: songs from same artist, 'Cold Fact' and 'Coming from Reality' from Sixto Rodriguez

#### d. Analysis output sample:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
27856	(0788806270, 078881172X, 6302787068)	(6302526574, 0788812807, 0788812408, 0788802194)	0.000313	0.000313	0.000313	1.000000	3190.000000	0.000313	inf
35798	(6303314015, 078881172X, 6302787068)	(6302158095, 6302526574, 0788812408, 0788802194)	0.000313	0.000313	0.000313	1.000000	3190.000000	0.000313	inf
35776	(6302158095, 6302526574, 6303314015, 0788812408)	(078881172X, 0788802194, 6302787068)	0.000313	0.000313	0.000313	1.000000	3190.000000	0.000313	inf

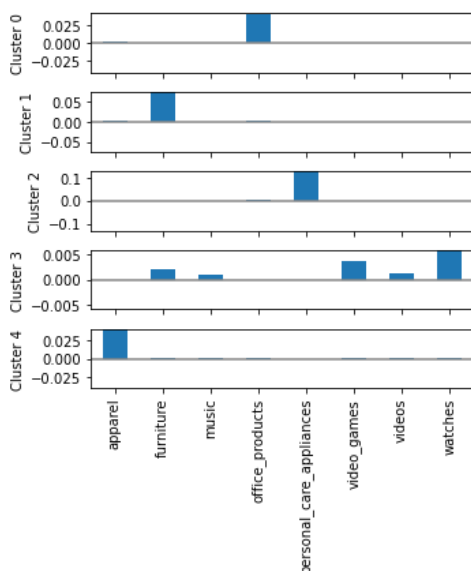
## 2. Customer segmentation based on product category | KMeans

- a. **Question:** Can we identify customer segments based on the purchased product categories to better target marketing campaigns?
- b. **Goal:** Help Amazon segment customers based on products bought by category. Target advertisements to specific customers within certain product segments in efforts to increase sales and revenues.

### c. Preliminary Insights:

- Product quantity bought not as relevant as product categories for customer segments
- 44% of customers already buying products from several categories
- 4 out of 5 customer segments (clusters) focus on only one specific product category
- 5 clusters emerged:
  - Cluster 0 = office products customer segment (86,103 customers)
  - Cluster 1 = furniture customer segment (78,795 customers)
  - Cluster 2 = personal care appliances customer segment (61,315 customers)
  - Cluster 3 = multi-category products customer segment (246,353 customers)
  - Cluster 4 = apparel customer segment (84,934 customers)

### d. Analysis output sample:



### 3. Topic analysis for customer reviews | NLP, LDA model

- a. **Question:** Can we extract key topics within product reviews to help companies analyze and interpret customer feedback?
- b. **Goal:** Help companies easily and readily extract key topics within customer product reviews. Identify positive or negative trends to maintain and/or improve their products and customer service.

#### c. **Preliminary Insights:**

- By dividing the analysis in good reviews (5 stars) and bad reviews (4 stars) we can infer the different adjectives emerging from the topics
- Example of salient terms
  - Bad reviews: bad, last, disappointed, slow, negative, terrible, comfortable, cheap, useless
  - Good reviews: great, good, easy, nice, happy, excellent, perfect, awesome, spare

#### d. **Analysis output sample:**

