QVC Business Analysis

Analysis and recommendations for enhancing marketing efforts and boosting sales

Prepared by:

Rachael Buchner | Emily Coppa | Libby Dondero | Lena Lin | Matthew Merrick

Today's Agenda

- Business problem
- Data preparation process
- Modeling techniques
- Cost-benefit analysis
- Descriptive statistics and clustering
- Recommendations
- Appendices

Key Results

Created a model with 91.35% accuracy

Segmented one-time purchasers into 4 customer segments

Quantified the success of advertising intervention as a net benefit of $49 Million

Action plan to solve the business problem

Rationale for Addressing Problem

- **Goal**: find which variables contribute the most to high-sales transactions
- These results will inform QVC's marketing strategy to increase sales
- Discover a model to decipher whether and how customers will generate high or low sales

Process and Considerations

- Grouped data into high and low classification → assumed 85/15% split between high and low sales
- Success criteria: 80% accuracy in our model
- Main potential costs: wasted airtime on viewers that do not convert to paying customers
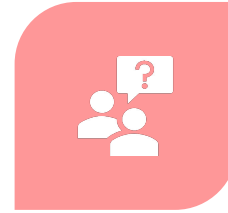
PROBLEM:
How to increase high-sales transactions

SOLUTION:
Find the key attributes and the most profitable customers

Key factors employed through data preparation step

Data quality issues

Missing values
Redundant values

Cleaning Techniques

Replace missing values
Remove useless attributes
Remove correlated attributes

Balanced Sampling

1431 High, 2500 Low

Feature Weighting

Picked top attributes

Split Data

80/20

After running 103 trials, a SVM model was most accurate

no.	Model	Accuracy	High Recall	Low Recall
1	**SVM**	**91.35%**	**91.96%**	**91%**
2	Voting	91.22%	84.62%	95%
3	Nearest Neighbor	90.97%	86.36%	93.60%
4	Voting	90.97%	87.06%	93.20%
5	Voting	90.97%	80.42%	97%
6	SVM	90.59%	86.36%	93.60%
7	Stacking	90.59%	86.36%	93%
8	SVM	90.33%	85.31%	93.20%
9	SVM	90.33%	87.41%	92%
10	SVM	90.20%	83.92%	93.8%

5 critical factors identified in the analysis

Transaction revenue: $72.33
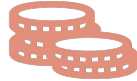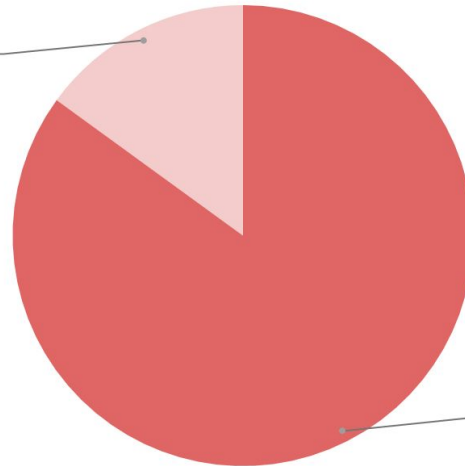
Advertising Costs: $0.05

Effective rate: 20%

Number of viewers: 24.6 million

Breakdown of the rate in sales

High sales
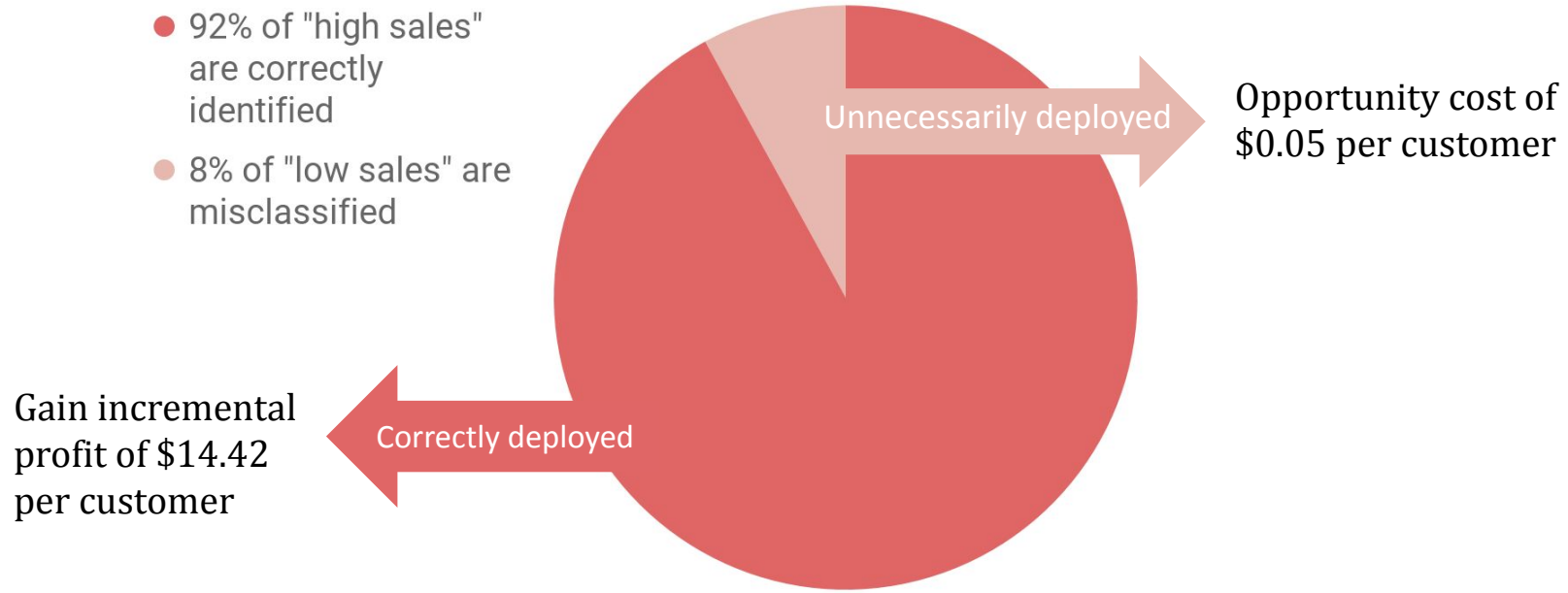*15.0%*

Low sales
85.0%

Overview | Predictive Analytics | Cost-Benefit Analysis | Descriptive Analytics | Recommendations

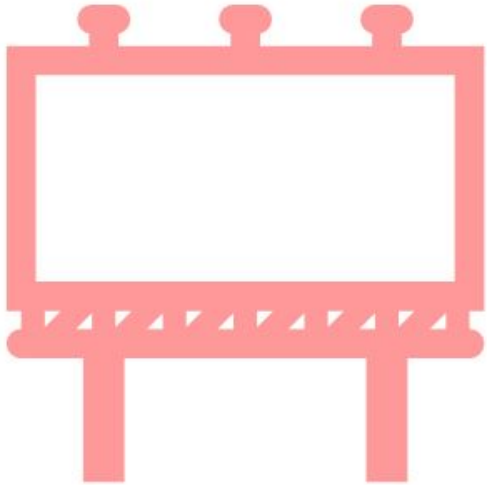# QVC can gain net benefit of $49 million by combining our model with targeted marketing

- 92% of "high sales" are correctly identified
- 8% of "low sales" are misclassified

Unnecessarily deployed

Opportunity cost of $0.05 per customer

Correctly deployed

Gain incremental profit of $14.42 per customer
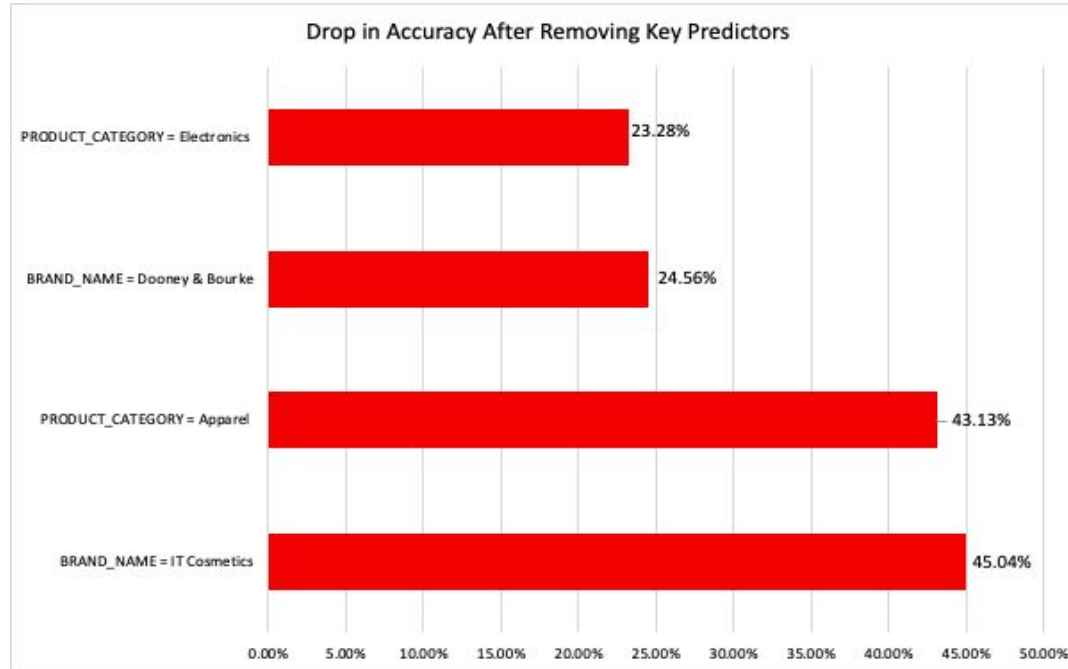
Key attributes from overall data

- Three key variables: "STATE", PRODUCT_CATEGORY", and "ORDER_PLATFORM"

- High sales generated from **New York**, with purchases of **home decor** products, and by ordering via **phone calls** after watching on-air segments
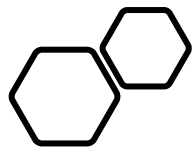
Key predictors of high sales from our model

Drop in Accuracy After Removing Key Predictors

- PRODUCT_CATEGORY = Electronics — 23.28%
- BRAND_NAME = Dooney & Bourke — 24.56%
- PRODUCT_CATEGORY = Apparel — 43.13%
- BRAND_NAME = IT Cosmetics — 45.04%

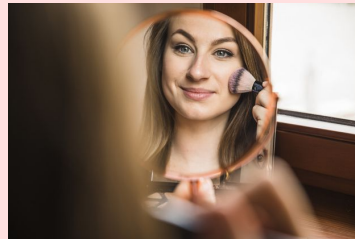# Clustering reveals four distinct consumer segments

Technological Tim

- Lives in NY
- Purchases electronics only
- Inquires about product details and all relevant information

Decorative Diane

- Loves home decor
- Seeks stylish and attractive products
- Wants to create a pleasing aesthetic

Trendy Tiffany

- Largest segment
- Active, trendy lifestyle
- Favorite brands include IT cosmetics, KitchenAid, and Aeropilates

Accessorized Amy

- High-end consumer
- Only purchases accessories like handbags
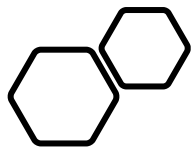- Prefers luxurious, high-quality products

Overview | Predictive Analytics | Cost-Benefit Analysis | Descriptive Analytics | Recommendations

Next steps for potential marketing strategies

Modeling
- Re-run Model 1 with a more conservative approach
- Conduct A/B testing

Marketing
- Focus efforts on predicted high-sales customers
- Promote electronic products
  - Modernize videography experience
  - Incorporate a Q&A segment
  - Create technology bundles
- Target "Trendy Tiffany"
  - Use a trendy, sophisticated presenter
  - Offer discounts to promoters of "Trendy Tiffany" products

Recommendations based on descriptive statistics

Marketing strategies based on clustering

Final Conclusions

Use optimal SVM model

$49 million net benefit

Focus on NY, home decor, and phone orders

Promote electronic products

4 customer personas: target "Trendy Tiffany"

Thank you!

Any questions?

Appendix 1

Data Preparation Process

Appendix 2

SVM Process

Parameters:

- Kernel Type: Anova
- Kernel Gamma: 5
- Kernel Degree: 2
- Kernel Cache: 200
- C: 0.1
- Convergence Epsilon: 0.001
- Max Iterations: 100,000
- Scale: Yes
- L Pos: 1
- L Neg: 1
- Epsilon: 0
- Epsilon Plus: 0
- Epsilon Minus: 0
- Balance Cost: Yes
- Quadratic Loss Pos: No
- Quadratic Loss Neg: No
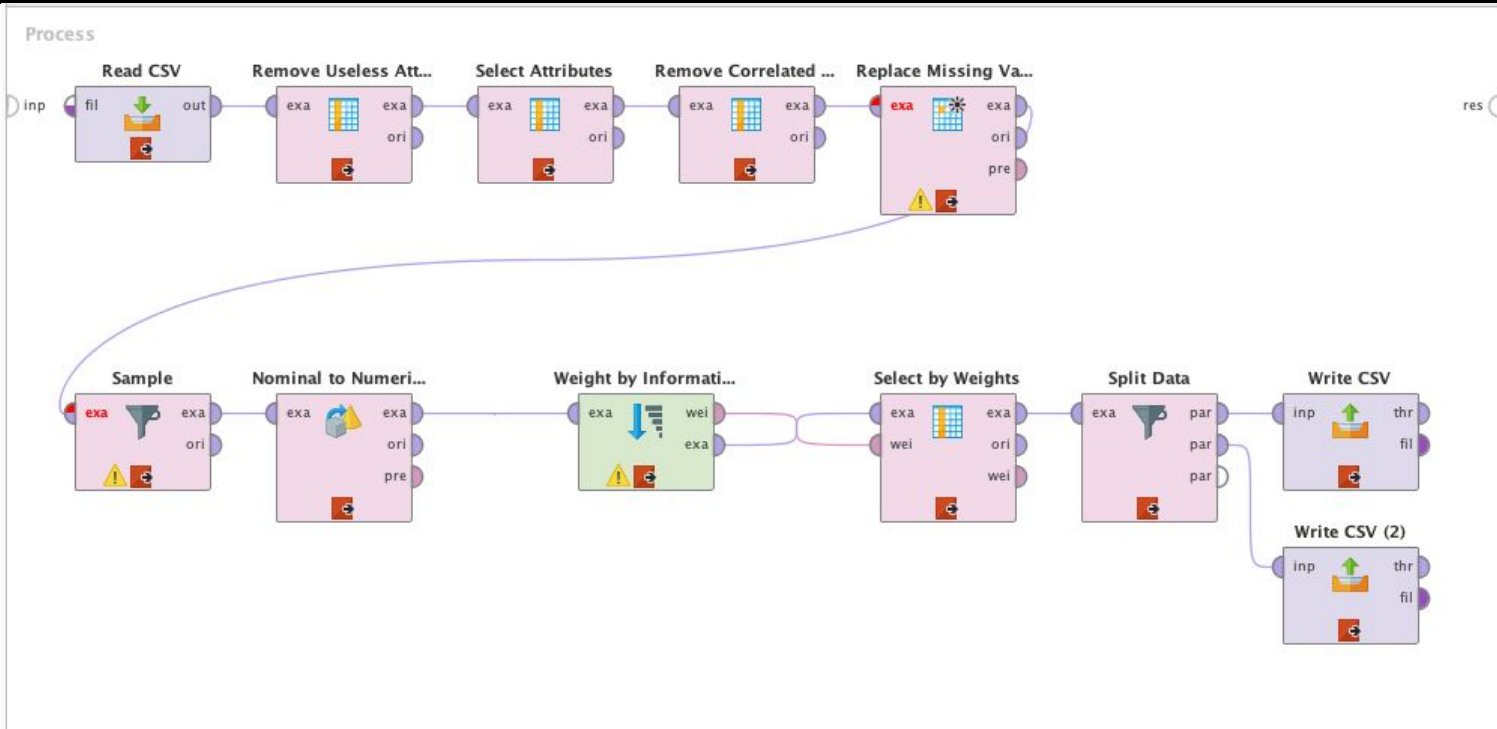
Appendix 3

Confusion and Cost Matrices for Optimal SVM Model

Confusion	Predicted Class		
Actual Class		**Class = High**	**Class = Low**
	Class = High	3,408,425.21	297,996.29
	Class = Low	1,881,414.99	19,023,196.04

Cost	Predicted Class		
Actual Class		**Class = High**	**Class = Low**
	Class = High	$14.42	$0
	Class = Low	-$0.05	$0

Appendix 4

Confusion & Cost Matrix Formula and Example Calculation

$$24.62m \; (number \; of \; viewers) = 5 \; \frac{million \; ad \; viewers}{year} \times 40 \; prime \; time \; ads \times .1231 \; \frac{qvc \; revenue}{walmart \; revenue}$$

True High Sales → $3,408,425.21 = 24.6m$ *(N)* $\times 15.06\%$ *(High Rate)* $\times 91.96\%$ *(High Recall)*

True Low Sales → $19,023,196.04 = 24.6m$ *(N)* $\times 84.94\%$ *(Low Rate)* $\times 91\%$ *(Low Recall)*

False High Sales → $1,881,414.99 = 24.6m$ *(N)* $\times 84.94\%$ *(Low Rate)* $\times (1\text{-}91\%)$ *(1-Low Recall)*

False Low Sales → $297,996.29 = 24.6m$ *(N)* $\times 15.06\%$ *(High Rate)* $\times (1\text{-}91.96\%)$ *(1-High Recall)*

True High Sales → $\$14.42 = [\$72.33$ *(transaction revenue)* $\times .20$ *(effective rate)*$] - \$0.05$ *(advertising costs)*

False High Sales → $\$(0.05)$ *(advertising costs)*

Appendix 5

Net Benefit Calculation

$Benefit = [transaction\ revenue \times effective\ rate \times true\ positive\ (high\ sales)] - advertising\ costs$

$Costs = false\ positive \times advertising\ costs$

$Benefit \rightarrow \$49,135,857.83 = 3,408,425.21 \times \$14.42$

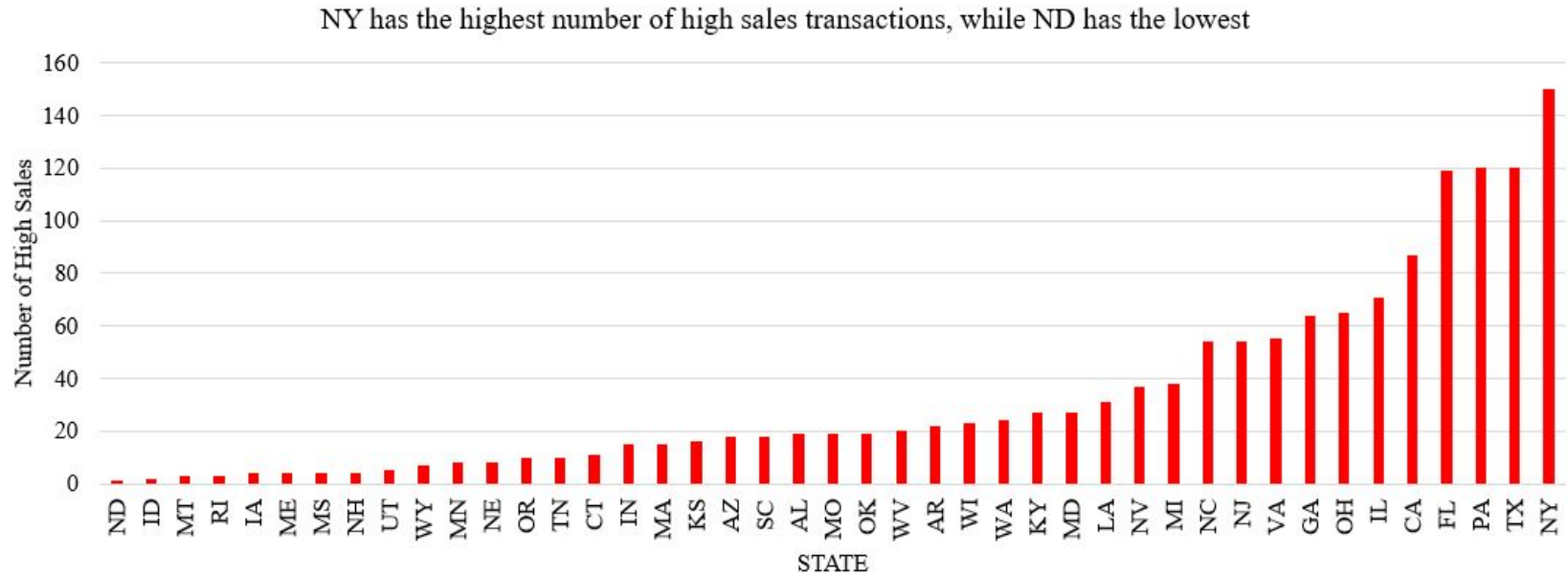$Cost \rightarrow \$(94,070.75) = 1,881,414,99 \times \$(0.05)$

$Net\ Benefit \rightarrow \$49,041,787.08 = \$49,135,857.83 + \$(94,070.75)$

Assumptions:
- N = 24.6 million
- High Rate = 15.06%
- Low Rate = 84.94%
- Effective Rate = 20%
- Advertising Cost per Impression = $0.05

Appendix 6

Descriptive Analysis: State

NY has the highest number of high sales transactions, while ND has the lowest

Appendix 7

Descriptive Analysis: Product Category

Home Decor products generate the highest number of sales, while IQVC Divisional products generate the lowest

Appendix 8

High sales transactions are mostly ordered through on air channels
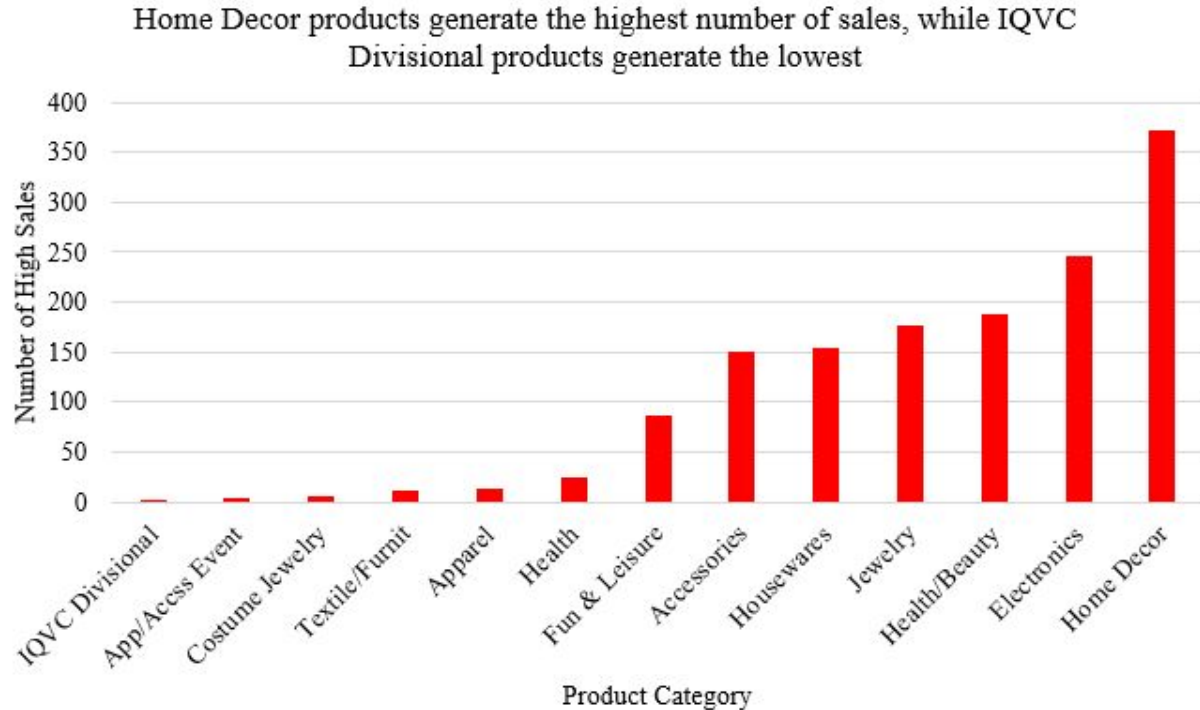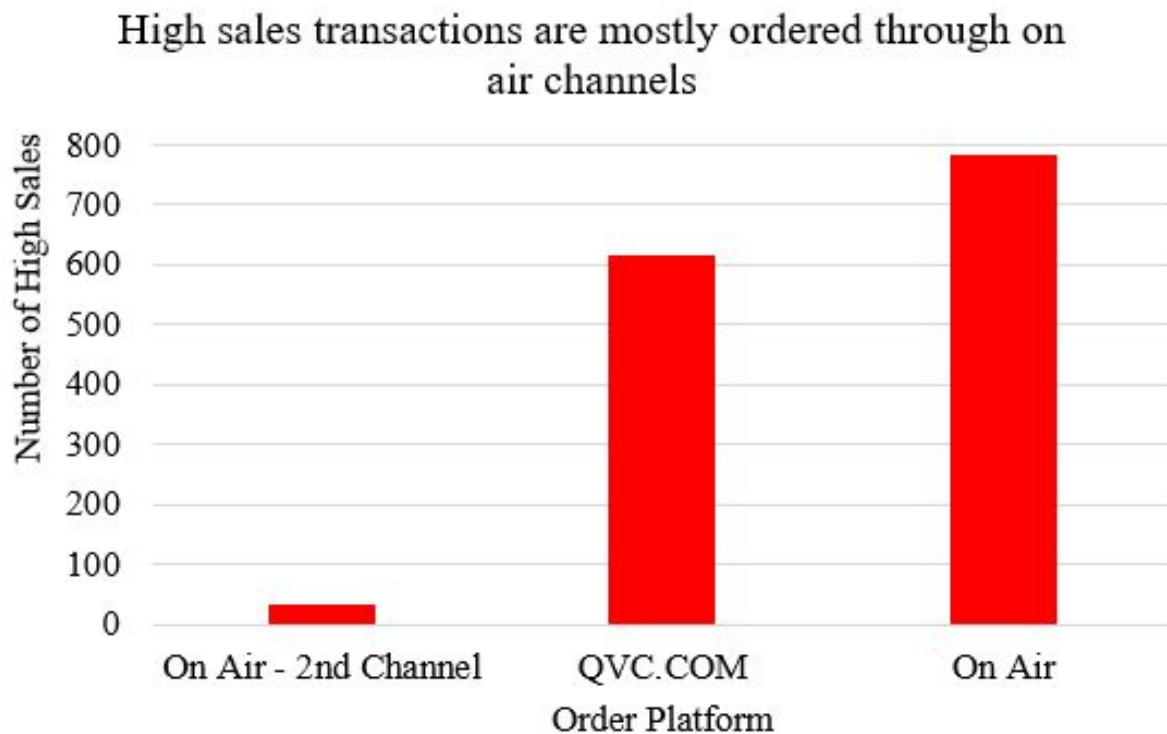
Appendix 9

Cluster Centroid Table

Attribute	Technology Tim		Attribute	Trendy Tiffany
PRODUCT_CATEGORY = Electronics	1		PRODUCT_CATEGORY = Home Decor	0.301
STATE = NY	0.355		PRODUCT_CATEGORY = Health/Beauty	0.182
STATE = FL	0.116		PRODUCT_CATEGORY = Jewelry	0.155
ON_AIR_MINS	10.905 mins		BRAND_NAME = Lenox	0.106
Share of Total	**11.60%**		BRAND_NAME = KitchenAid	0.065
			BRAND_NAME = Dyson	0.044
Attribute	**Decorative Diane**		STATE = NY	0.09
PRODUCT_CATEGORY = Home Decor	0.404		STATE = FL	0.085
PRODUCT_CATEGORY = Jewelry	0.163		ON_AIR_MINS	6.38 mins
PRODUCT_CATEGORY = Housewares	0.154		**Share of Total**	**56.18%**
PRODUCT_CATEGORY = Health/Beauty	0.128			
STATE = FL	0.093		**Attribute**	**Accesorized Amy**
ON_AIR_MINS	7.633 mins		PRODUCT_CATEGORY = Accessories	1
Share of Total	**21.80%**		Brand_Name = Dooney & Bourke	0.805
			Brand_Name = G.I.L.I.	0.101
			STATE = CA	0.141
			STATE = CT	0.048
			ON_AIR_MINS	2.994
			Share of Total	**10.41%**

Appendix 10

Cluster Size

Cluster Model

```
Cluster 0: 166 items
Cluster 1: 312 items
Cluster 2: 804 items
Cluster 3: 149 items
Total number of items: 1431
```