The Customer Is Always Right: Analyzing Existing Market Feedback to Improve TVs

Jose Valderrama[1], Laurel Rawley[2], Simon Smith[3], Mark Whiting[4]
[1]University of Central Florida
[2]University of Houston
[3]University of Wisconsin-Madison
[4]Carnegie Mellon University

Abstract
Online consumer reviews can be analyzed using an algorithm that quantifies the consumer's sentiments towards a product as well as the sentiment towards specific features of a product. In turn, the covariance between different features can be analyzed and rated. Our research uses both feature and sentiment analysis to illustrate these correlations and consumer preferences.
Keywords: sentiment; features; products; covariance; consumers
doi: 10.9776/16589

Acknowledgements: Thanks to Mike Depew and everyone at the i3 program
Contact: josevalderrama@knights.ucf.edu

1 Introduction

In the last few years, user generated content has rapidly increased through the means of social media and product review sites. A continuous stream of feedback is accessible to consumers. In addition, companies are able to use these reviews in order to monitor customer satisfaction. We analyzed product reviews from Amazon.com that referenced the opinions and experiences of a customer towards the given product, in this case television sets. Using a sentiment analysis of the overall satisfaction of users through textual mining, consumers and companies can conclude valuable interpretations of their products' features.

2 Background

We found several related works that analyze features and products. In Eirinaki, Pisal, and Singh (2012), the authors explore feature-based opinions and how to rank them. The authors note 67% of internet users regularly use social network sites and/or blogs (Eirinaki et al., 2012). With the increase in internet usage, "most people are using the Internet to check the reviews of products before buying them" (Eirinaki et al., 2012), The algorithm proposed in this paper first determines if a review is positive, neutral or negative and then it continues on to identify features and their corresponding reviews.

Additionally, in Linden, Smith, and York (2003), an overlook of Amazon recommendations is given. The four algorithms used in Amazon are collaborative filtering, cluster models, search based, and item-to-item filtering. Collaborative filtering involves a customer being compared to other customers to use previous patterns to predict a customer's preference in products. The cluster models place each customer in a segment, and recommendations within that grouping are given to the customer. Search based recommendations are based on keywords and similar subjects. Lastly, item-to-item filtering recommends products that are frequently bought together. This is especially useful for large data bases and matches items with other similar items. Previous research used several methods in order to gather data which include the compilation of consumer review text and the division of the text into feature-based segments to determine whether a review is positive or negative (Archak, Ghose, & Ipeirotis, 2011). Popescu and Etzioni (2005) describe an automated information extraction system called OPINE, which can be used to automatically perform a review analysis.

3 Methods

The methods in this paper consist of computational analysis of a collection of raw data pulled from Amazon through the use of extraction tools from import.io. Import.io is an accessible program that extracts textual data and sorts it into table formats, specifically excel files. The reviews (each of which consists of a numerical five-star system and a written review) were analyzed by pinpointing keywords, i.e. feature names, within the review in order to compute an overall sentiment of a product and its features. We extracted nearly 1,000 reviews, 40-60 reviews from each of our selected 25 television sets. Commodity computers were used as opposed to servers or mechanical turk as a result of our need for identical formats from the collected reviews.

The data collected consists of reviews of television sets ranging from 30-43 inches with prices under $500. This category represented the largest collection within the television market on Amazon. This provides us with a substantial amount of useful data that can help us conduct an accurate sentiment analysis on our collection of reviews and their characteristics.

Our feature extraction method involves identification of key words, or vocabulary, and assessing the effect of that word and the overall sentiment towards a product. Features were chosen based off of product descriptions located on Amazon. Sentiment analysis for our reviews was done using TextBlob ("TextBlob", 2015) and the Natural Language Toolkit for Python (hereafter referred to as NLTK) ("Natural Language Toolkit", 2015). Review data, originally extracted in csv format, was parsed and tokenized using NLTK. Reviews were grouped based on their corresponding product, giving us one group of reviews for each product. Our method of using TextBlob performs sentiment analysis on individual sentences. Each sentence is assigned a polarity score between -1.0 and 1.0, which is then adjusted to match the star rating scale of 0 to 5; by averaging out these scores of an entire review, an overall sentiment can be generated for the review. These review sentiment scores can be further averaged across all reviews of a single product, giving an overall sentiment for each product. The average of all star ratings that originally accompanied the review text is also obtained for each product. An idea of the consumer opinion of a product can be obtained by comparing both the average star rating and average sentiment rating of a product.

4 Results and Analysis

Generally, the behavior of the data analyzed in this report was accurately predicted using theoretical statistics. With assumptions, such as treating the measured values as random variables, the sample data could then be modeled efficiently. The importance, as well as overall satisfaction, of the feature to each consumer played a role in its score. Most reviews mentioned screen size along with pricing when critiquing the given product which shows that they are the two most popular features within the sample data. The covariance between two television sets based on their ratings was also used in order to model any dependencies between features and whether the effectiveness of one feature played a role in the customer's satisfaction with another. Figure 1 shows a table containing the sentiment analysis scores of every feature for every product. Blank entries indicate the feature was not mentioned for that product, and there is no data. Highlights in the table indicate the lowest score that occurred for that feature.

We provide heat maps in Figures 2 and 3 of our sentiment analysis results. This displays the sentiment scores of every feature in every product as a matrix, along with additional data for average star rating and overall sentiment score for each product, and average feature sentiment for each feature. Darker red colors indicate a less favorable response, while brighter yellow or white colors indicate a more favorable response. Orange colors indicate an average response, while black sections of the heat map indicate no data (the feature was not mentioned for that product).

At the screen size with the most competitors, the favorability of the screen is inversely proportional to the star rating. As the screen size increases, this relationship seems to flip which reflects the influence of the reviews including screen size. At this screen size, the higher star rating coincides with the higher sentiment score for the screen. Generally, consumers are neutral about brands but seem to express preferences for brands that have a more favorable life cycle. This could be because consumers rely on these products lasting longer and functioning properly. Therefore, they are more inclined to buy products that they know will work longer to get the most of the money spent on a product.

Product List Sorted by Star Rating

TV Name	Description	Average Star Rating	Unit Price	Overall Sentiment	Price	setup	HD	conn	smart	3D	settin	brand	hard	resol	aspec	HDMI	life	refres	year	LED	rating	mode	scree	scree	audio	warra	ratio	screen
LG ELECTRONICS 39	39-Inch 1080p 60Hz LED	4.75	320.00	2.96	3.17	3.35	2.74		2.89	2.76	3.03	2.82		3.06		2.80	2.82	2.68	3.14	3.16	3.38	2.49	3.14	0.36	2.75	3.24		39
Seiki SE39UY04	39-Inch 4K Ultra HD 120H	4.60	474.47	2.78	3.21		2.90				3.00	2.32				2.93	2.10		2.19			2.92	2.70	-0.09	2.84	2.56		40
Upstar P40EC6	40-Inch 1080p 60Hz LED	4.54	279.00	2.97	3.24	3.46	3.32			3.19	3.14	3.05		1.25	4.50	2.78	2.50	3.35	2.97	2.31	2.85	2.83	3.01	0.22	2.58	2.47	2.46	42
Panasonic TC-L42U25	42-Inch 1080p 120 Hz LC	4.52	469.59	2.84	3.24	2.94	3.03		2.46	2.65	2.67	2.67	2.63	3.08	2.50	2.58	2.50	2.71	2.69	2.74	2.97	2.42	2.86	0.07	2.88	2.73	2.71	40
Toshiba 40L1400U	40-Inch 1080p 60Hz LED	4.50	346.75	2.95	3.25	3.09	3.05	2.98	2.98	2.96	2.93	2.90	2.90	2.87	2.85	2.84	2.83	2.83	2.82	2.81	2.79	2.79	2.79	0.00	2.77	2.71	2.68	40.32
Westinghouse UW40T	40-Inch 1080p 120HZ Sli	4.38	389.99	2.92	3.23	2.70	3.14				2.82	2.19		2.69	2.44	2.71	2.96		3.28		2.50	2.61	2.59	-0.20	3.00	3.24		42
oCOSMO	40-inch 1080p 60Hz Roku	4.33	299.99	2.90	3.40	2.54	3.44	2.94	2.47	2.92	2.90	2.91		2.75		2.91	2.50	2.58	2.84	3.84	2.50	3.23	2.81	0.02	3.18	2.52	2.58	40
LG Electronics 42LB5	42-Inch 1080p 60Hz LED	4.33	399.99	3.11	3.64				2.99		3.31	3.46	3.63			2.79			2.55	2.99	2.84	2.80	2.80	0.02	2.82	2.81		39
LG Electronics 42LB5	42-Inch 1080p 60Hz Sma	4.28	448.91	2.93	3.23	3.10	2.58		2.87	2.28	2.27	2.34	2.18	2.53		2.81			3.03	2.97	2.69	2.50	2.65	-0.13	2.63	2.58	2.05	40
Sharp LC-39LE551U	39-inch Aquos HD 1080p	4.28	299.00	3.19	2.50		2.67		4.02		3.75					2.30			3.69	0.75			3.00	0.21	3.06			40
Samsung UN40H5003	40-Inch 1080p 60Hz LED	4.27	324.74	3.01	3.15	3.38	3.11	3.50	3.13		3.28	2.47				2.95		2.52	2.70	3.36	2.50	2.49	2.74	-0.04	2.72	2.93	3.09	40
Toshiba 40E210	40-Inch 1080p LCD HDT	4.26	349.99	3.04	3.24	3.67	2.82		3.07	2.50	2.40	3.58				2.75		2.50	2.75	2.27		2.61	2.28	-0.50	2.50			40
Upstar P40EA8	40-Inch 1080p 60Hz LED	4.04	269.00	2.99	3.62	3.13	3.13	3.50	3.13	4.06	3.03	2.48		3.33		2.79		2.52	2.70	3.29	2.50	2.58	2.81	0.03	2.43	2.93	2.17	42
RCA LED42C45RQ	42-Inch 1080p 60Hz LED	4.03	293.29	2.98	3.14	3.03	2.62		2.80		3.01	2.94			2.77	2.68			2.82	2.82	2.54	2.94	2.87	0.08	2.90	2.61	2.77	40
Sony KDF-E42A10	42-Inch LCD Rear Projec	4.02	349.99	3.07	3.37	3.91	2.99				3.11	2.73		3.00		2.67			2.66	3.11		2.50			2.48	2.12		40
Samsung LN40A750	40-Inch 1080p DLNA LCD	3.97	429.00	2.76	3.07	2.92	3.28	2.50	3.05		2.51	2.77	2.92	2.78		3.22	2.59	3.45	2.48	3.26	2.58	2.71	2.84	0.05	2.35		3.06	40
Sony KDL40R380B	40-Inch 1080p 60Hz LED	3.70	397.72	3.17	3.76		4.06		2.56		3.21					2.59						3.34	2.55	-0.24				40
Seiki SE40FY27	40-Inch 1080p 60Hz LED	3.66	279.00	2.90	3.53	2.52	2.57		3.02		2.77	2.84		3.33		2.74	2.67		2.58	2.97		2.84	2.51	-0.27	3.08	2.62		40
Sceptre X409BV-FHD	39-Inch 1080p 60Hz LED	3.62	255.99	2.85	3.32	2.75	2.79		2.90	2.69	2.90	3.04			3.36	2.79	2.78	2.79	2.80	3.08	2.50	3.11	2.86	0.07	3.04	3.17	3.50	40
TCL LE40FHDE3010	40-Inch 1080p 60Hz LED	3.57	268.88	3.00	3.08	2.29	3.17	2.93		3.63	3.24	3.21		2.78		2.83			2.08	2.17	4.50	3.11	3.08	0.30	2.67	2.50		40
VIZIO E371VL	37-Inch Class LCD HDTV	3.32	395.00	3.12	3.38	2.89	3.80		3.03	3.04	3.23	2.90	3.31	3.50		2.86		2.77	2.92	2.82		3.49	3.11	0.32	3.25		2.53	42
VIZIO M422i-B1	42-Inch 1080p Smart LED	3.23	428.59	2.76	3.16	2.97	3.13		3.11		2.34	2.87	2.92	3.35		4.00	2.64	3.58	2.44	3.47	2.24	2.73	2.79	0.01	2.30			42
VIZIO E400i-B2	40-Inch 1080p Smart LED	2.88	379.00	3.01	3.44	3.56	2.91	2.50			2.63	5.00			2.50	2.97	4.69	2.45	3.65	2.76	2.50	2.50	2.78	-0.01	2.67	2.83	2.73	42
Coby TFTV4025	40-Inch 1080p 60Hz LCD	2.77	345.00	2.75	2.88	2.51	2.85		2.85	2.83	2.86	2.52	2.67	2.73	2.43	2.74		2.59	3.26	2.40	2.38	2.61	2.76	-0.03	2.73	2.81	2.69	39
VIZIO E420i-B0	42-Inch 1080p LED Smart	2.76	399.99	2.88	2.95	2.99	2.85		3.00	2.93	2.92	2.61	2.92	2.90	2.92	2.97	3.26	3.22	2.78	2.70	3.50	2.67	2.81	0.03	2.69	2.34	2.96	42

Figure 1. Individual Feature Scores

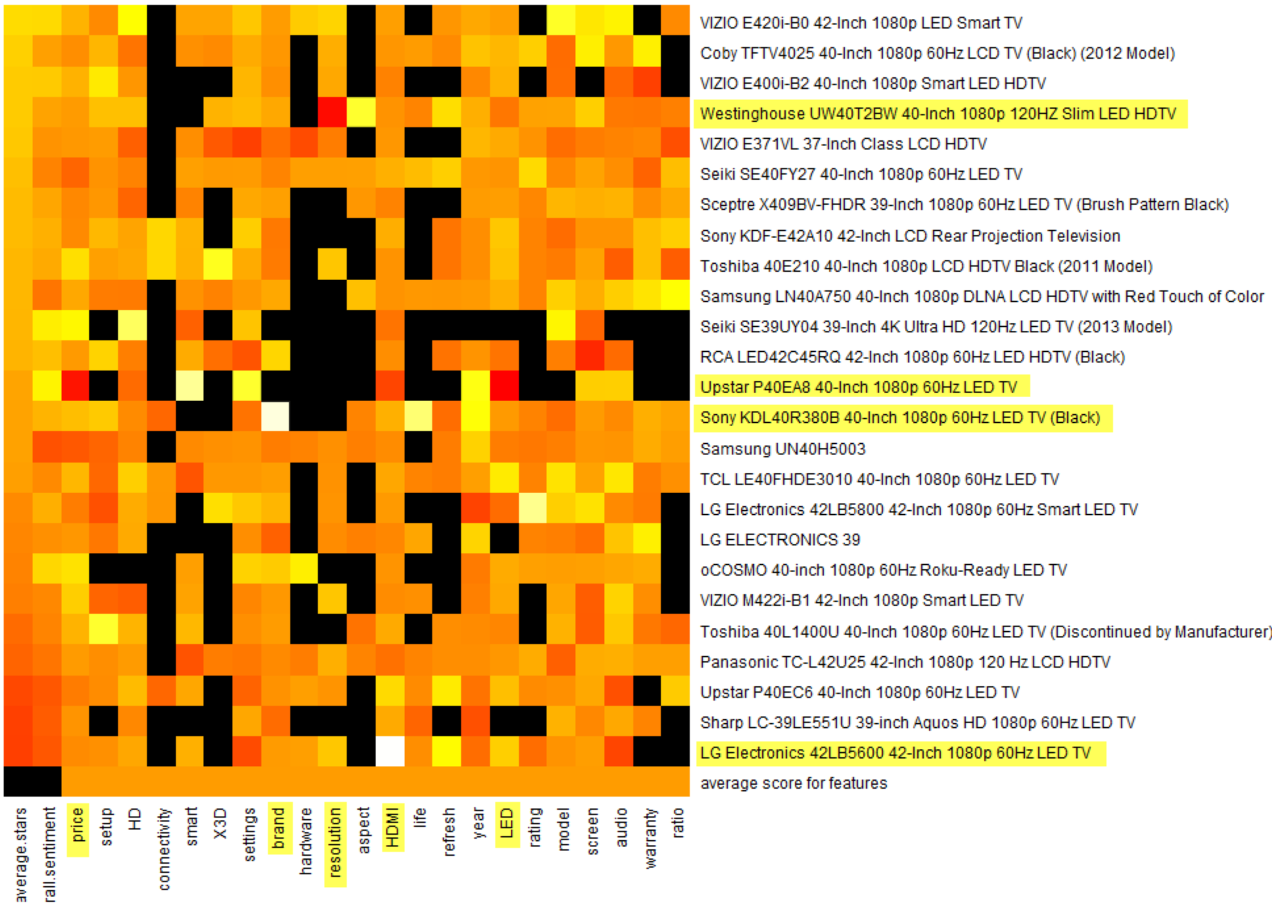

Figure 2. Feature Scores Heatmap; notable results highlighted

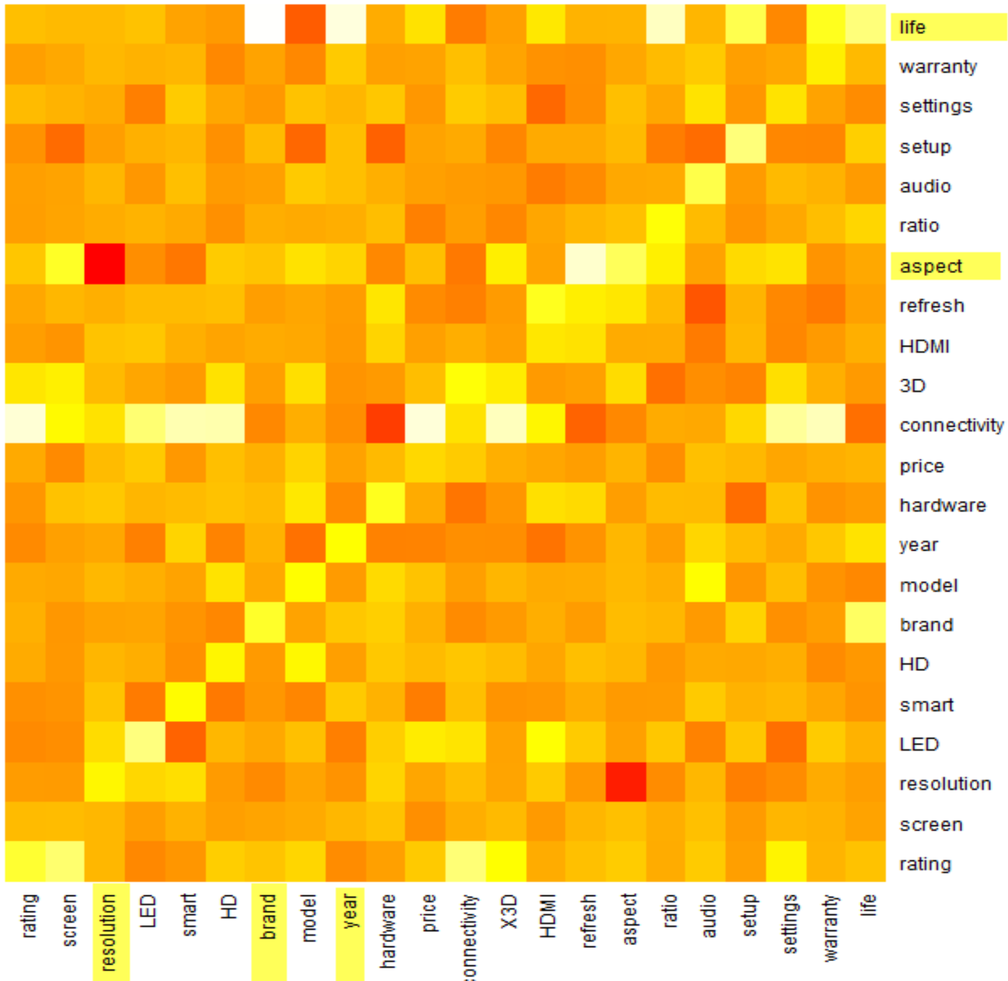

Figure 3. Covariance Among Features; notable results highlighted

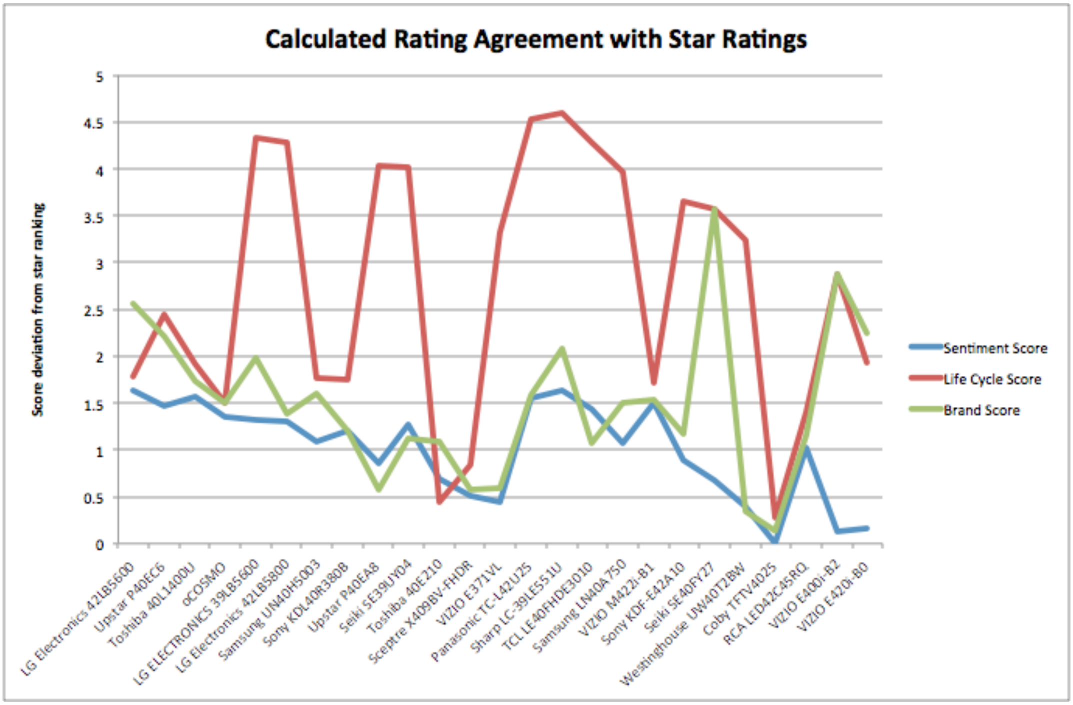

Figure 4. Calculated deviations from Star Rating using root mean squared

5 Conclusion

Our research can help industries optimize their products to maximize customer satisfaction. By using sentiment and feature analysis, it is evident that some characteristics of a product are much preferred than others. By creating a covariance matrix, shown in Figure 3, among television features, interesting relationships are illuminated. Aspect and ratio and resolution have a very strong negative relationship whereas life and brand have a very strong positive covariance. However, the number of television reviews that mentioned these features are not enough to warrant a general conclusion. Companies can take advantage of relationships such as these to improve functionality of features to which consumers are most responsive.

Figure 4 shows the agreement of sentiment and star rating as well as the two features with the highest covariance score. The behavior of the data implies that there is a direct correlation between the sentiment scores of 'life cycle; and 'brand' features. Overall, a product's single feature and its corresponding rating does not tell consumers much about the product. However, with more features and ratings for a product, a more holistic and accurate rating can be determined.

While our findings provide a novel means to optimize products and determine what it is people want, future research could include the incorporation of more statistical analysis as well as taking into account second mover advantage in research. However, first movers can also use the information found to optimize their current products to stay ahead. In the end, companies will be competing for consumers' preferences.

In addition, a larger data set may be analyzed, whether it be on televisions or other consumer goods. Analyzing more data can further illuminate precisely what consumers want and how companies can adapt to fulfill consumer needs. The larger the data set, the more accurate the results will be. More complex models could construct an optimal set of features for best possible market performance. Companies can use the data to optimize pre-existing or future products to improve sales and customer satisfaction.

6 References

Archak, N., Ghose, A., & Ipeirotis, P. (2011). Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science 57(8)*:1485-1509. Retrieved from http://dx.doi.org/10.1287/mnsc.1110.1370

Eirinaki, M., Pisal, S., & Singh, J. (2012). Feature-based opinion mining and ranking. *Journal of Computer and System Sciences, 78(4)*, 1175–1184. Retrieved from http://ac.els-cdn.com/S0022000011001139/1-s2.0-S0022000011001139-main.pdf?_tid=faf57c32-24e6-11e5-9d65-00000aab0f6b&acdnat=1436301197_d10e1ce3d45e82d736decf75c001c63a

Linden, G., Smith, B., & York, J. (2003). Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing, Jan.-Feb. 2003*. Retrieved from http://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf

Natural Language Toolkit. (2015, September 5). Retrieved from http://www.nltk.org/

Popescu, A., & Etzioni, O. (2005). Extracting Product Features and Opinions from Reviews. *HLT '05 Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing:* 339-346.

TextBlob: Simplified Text Processing. (2015). Retrieved from http://textblob.readthedocs.org/en/dev/

The Second-Mover Advantage - A Primer on How Late-entering Companies Can Compete with Pioneers. (2013, November 4). Retrieved from http://insight.kellogg.northwestern.edu/article/the_second_mover_advantage