Judy Huang, Nicole Arugay, Evan McClane, Angela Sherpa
MKT4561 Project 3 Analysis

1) Introduction

Our group chose [a dataset from Amazon Alexa] that recorded data on customer reviews and ratings for Alexa Echo and Echo Dot.

2) Marketing Questions

By taking into account the different data that we have, we want to answer our main question "Are customers satisfied with their purchases?" We will further dissect this question by understanding the factors that contribute to customer sentiment.

1. What function are Alexa products mainly being used for?
2. Do product ratings correlate to the sentiment of customer reviews?
3. What is the overall sentiment of Alexa products?

3) Data Cleanup & Coding

Stemming & Stopping																																	
Document 1	Echo	sound	same	emoji	speaker	from	Five	Belome	sibling	has	even	one	has	Bluetooth	not	need	plug	in	only	good	thing	about	Echo	you	can	talk	Echo							
Document 2	Echo	respond	when	not	talk	Echo	unplug	Echo	Echo	feel	like	Echo	spy																				
Document 3	not	like	almost	everytimeme		ask	Echo	question	Echo	would	talk	Echo	not	know	or	Echo	has	not	learn														
Document 4	you	think	about	really	not	do	much	play	music	answer	stupid	question	finally	Echo	able	hook	thermostame		can	see	why	price	drop	like	rock									
Document 5	not	all	happy	speaker	not	great	reason	lot	time	when	play	music	Echo	will	start	make	weird	noise	stop	child	has	Google Ho	much	better										
Document 6	sad	joke	worthless																															
Document 7	start	talk	Echo	self	has	not	use	Echo	since	unplug	Echo	day	not	one	talk	child	was	lay	bed	next	Echo	just	start	go	about	movie	Moana	completel	scare	sh*t	out	child	now
Document 8	buy	Echo	healing r	play	continuou	spa	music	although	Echo	know	type	music	need	Echo	not	deliver	quality	sound	Echo	sound	muffle	not	know	buy	speaker	too							
Document 9	love	Echo	dislike	can	not	sync	Apple Mus	Echo	play																									
Document 10	primarily	music	bedroom	wish	speaker	better	fidelity																											
Document 11	Echo	not	bad	has	issue	Echo	actually	play	want	get	Echo	use	iHeartRad	properly	play	station	want	challenge	Echo	not	bad	sale	price		paid								
Document 12	child	love	Echo	skill	amazing	like	see	Echo	get	little	smarter	answer	basic	question																				
Document 13	love	new	Echo	now	mainly	use	Echo	play	music	night	Echo	get	job	done																				
Document 14	has	three	Echo	now	even	one	barn	great	idea	when	work	horse	still	get	news	flash	music																	
Document 15	spoil	Echo	downstair	get	Echo	upstairs	listen	music	etc	when	walk	treadmill																						

Cleaned Data

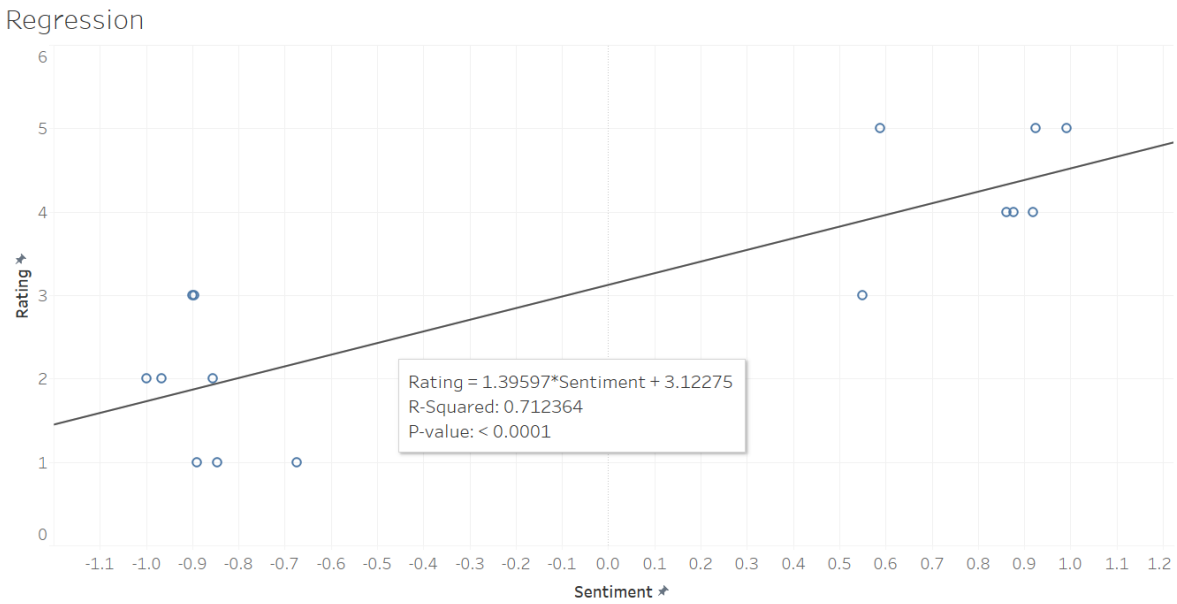
We cleaned the data by first taking a random sample of three reviews at each rating level. We then used stemming to match words to their root form and eliminated stop words to evaluate documents based on the important words. This eliminates the need to evaluate words such as "the," "is," "a," and so on that are not significant in our analysis. Finally, we used tf-idf coding to evaluate the relevance of each word in the corpus.

4) Analysis by Section

Voyant Word Cloud

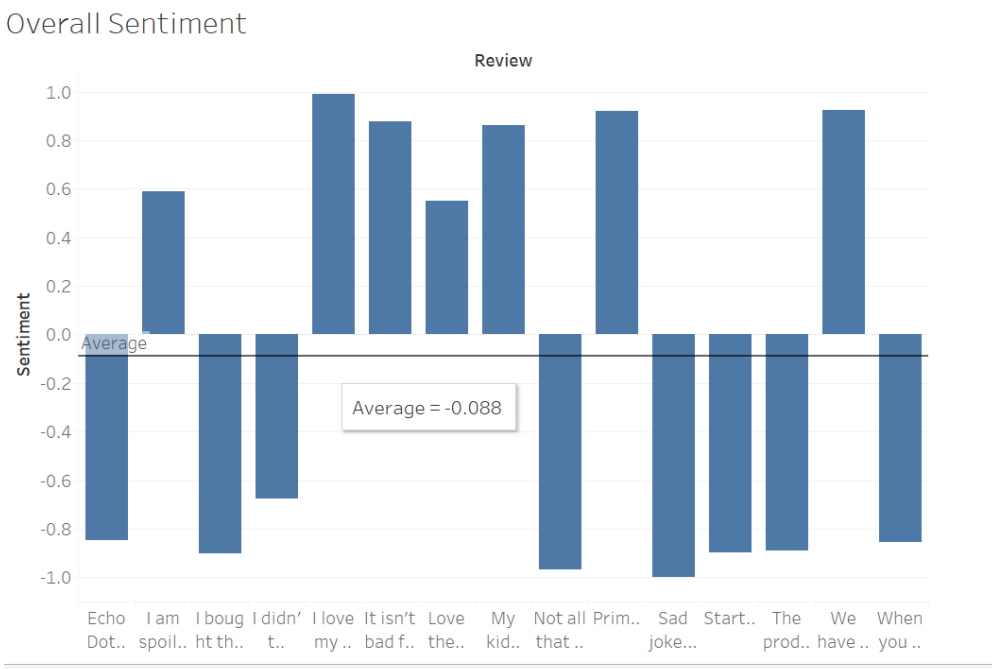
Our first subquestion addresses what function Alexa products are mainly used for. We used Voyant Tools to create a word cloud based off of customer reviews to see what customers typically use Alexa for. We see words like 'music', 'play', 'speaker' are in the biggest font. From this, we can conclude that Alexa

products are used mainly for playing music. We know that this is the main function that customers are looking for so customer sentiment will be based on how well Alexa products perform this function. Some other uses include asking questions and Alexa replying.



Rating vs Sentiment Regression Model

Our second subquestion addresses whether product ratings correlate to the sentiment of customer reviews. We used the Monkey Learn sentiment analysis tool to get the sentiment of each review. We then created a linear regression model on Tableau to compare product ratings to the sentiment of reviews and got the equation y = 3.12 + 1.4*(Sentiment) indicating a positive correlation between rating and sentiment. The R^2 value is 71% and p-value of <.0001 shows that there is a strong positive relationship as 71% of the ratings can be explained by sentiment. This tells us that ratings are a great way to get a general sense of what customer sentiment is like. We do have to, however, take into account some outliers from people that are more harsh or lenient in their ratings compared to what their reviews say. For example, a customer gave a 2 star rating with a review that said “Sad joke. Worthless.” leading to a -100% sentiment. In this case, the customer was too lenient with their rating and should’ve rated it 1 stars given what they said in their review.



This corpus has 1 document with 415 total words and 218 unique word forms. Created now.

Vocabulary Density: 0.525

Average Words Per Sentence: 14.3

Most **frequent words** in the corpus: music (9); like (5); play (5); alexa (4); love (3)

Overall Sentiment Bar Graph and Average Line

Voyant Tools Most Frequent Words

Our third question addresses the overall sentiment of Alexa products. We used the sentiments that we gathered from Monkey Learn to create a bar chart that shows the sentiment of each review. I included an average line that indicates an overall sentiment of -0.088 or -9%. From the graph, you can see that positive and negative reviews are almost even with 7 positive reviews and 8 negative reviews. Due to there being 1 more negative review, the overall sentiment is a weak negative of -9%. We can conclude that the overall sentiment towards Alexa products is generally neutral with customers having both positive and negative experiences. We also looked at the most frequent words in the corpus. Interestingly, Voyant Tools indicated that words like 'like' and 'love' were the most frequent words but it's important to note that surrounding words can change their meaning entirely. For example, 'like' can be used in the context of 'feels like', 'spa like music', and 'didn't like' which shows that most frequent words may not be the best indication of what the overall sentiment towards Alexa products are.

Row No.	prediction(Sentiment)	confidence(negative) ↑	confidence(positive)	text
1	positive	0.443	0.557	product sounded ...

Sentiment Analysis using RapidMiner

Using RapidMiner's built-in Sentiment Analysis algorithm, we see that the model used predicted the overall sentiment to be positive, with a negative confidence of 0.443 and a positive confidence of 0.557. Because the model predicted almost an even amount of confidence levels for a negative and positive review, we can see that the model used generally reflects our findings so far which is that the sample of

reviews examined can be described as more neutral. The algorithm used in this built-in sentiment analysis model is as follows:

1) Importing historical sentiments with some assessment such as extracting the words and delivering a word-vector
2) Train a support-vector machine model and validate it to collect new performance data
3) Create a new document from text and processing it as the initial ones
4) The model trained with the old text is applied to the new document

Step three is when we uploaded the new text document containing Alexa's reviews which was when the model was applied.

5) **Conclusions**

To answer our main question of whether customers are satisfied with their purchases, we looked at what function customers are most concerned about, whether product ratings correlate with customer review sentiment, and what the overall sentiment towards Alexa products were. We concluded that customers are mainly concerned about the music playing function and have a generally neutral sentiment towards the products. Customer experience varies greatly with some having -100% sentiment and others having +99.2% sentiment towards the products. We also learned that ratings are a great way to get a general idea of what the customer's sentiment towards the product is, while many frequent words lack context and are not an effective way of showing satisfaction on their own.

We also found that there were nuances to look out for when analyzing sentiment. As said previously, we have to make sure the model accounts for different uses of words and to look out for phrases. For example, the word "like" can be used in various contexts that can be both positive and negative. Another factor to consider is user error. When analyzing sentiment, we should take into account any mistakes a user might make when writing their review or if anything is misaligned. The example given earlier was the arguable leniency of one user's star rating which misaligned with their written review that said "Sad joke. Worthless." In analyzing sentiment, we can also look out for user biases such as paid or sponsored user reviews.

6) **Recommendations**

From the reviews, it is clear that customers have drastically different experiences with the same product. A recommendation for Amazon is to look at the pain points of the product through customer reviews and fix them in the next product launch or upgrade so that instead of an overall sentiment of -9%, customers would have a positive sentiment towards Alexa products. According to our text analysis, we also found that music is one of the main use cases for these products.

Amazon can also simultaneously pay attention to both their 30-40 year old age demographic and young children demographic. According to the word cloud, one relevant word included "daughter" which could indicate that parents are leaving reviews and relying on their children's experiences to form their opinion. Amazon can look into how children can have a positive experience with their products. In general, curating a product that makes it possible to make memorable and long-lasting memories that are pleasant make for good retention rates for the product. Amazon can utilize this marketing strategy and use sentiment analysis to see what their consumers are interested in.

7) **Appendix - Screenshots**

Corpus	
Document 1	The product sounded the same as the emoji speaker from five below my sister has ... and even that one has Bluetooth and doesn't ne
Document 2	Echo Dot responds to us when we aren't even talking to it. I've unplugged it. It feels like it's spying on us.
Document 3	I didn't like that almost everytime i asked Alexa a question she would say i dont know that or i havent learned that.
Document 4	When you think about it this really doesn't do much. Play music, answer stupid questions, finally it was able to hook up to therm
Document 5	Not all that happy. The speaker isn't great and for some reason a lot of times when I play music it will start to make some weird
Document 6	Sad joke. Worthless.
Document 7	Started taking on her own! Haven't used it since I unplugged it that day. No one was talking, my daughter was laying in bed next t
Document 8	I bought this for my healing room. I play continuous spa like music. Although Alexa knows the type of music needed, the device do
Document 9	Love the product but dislike that I can't sync up my Apple Music for Alexa to play
Document 10	Primarily for music in our bedroom. I wish the speaker had better fidelity.
Document 11	It isn't bad for what it is. Have issues with it actually playing what I want. Getting it use iHeartRadio properly by playing the
Document 12	My kids love Alexa... Skills are amazing, would like to see her get a little smarter answering basic questions
Document 13	I love my new Echo Dot! Right now, I mainly use it to play music at night, but it gets the job done.
Document 14	We have 3 echos now, even one in the barn, great idea for when I am working with my horse, can still get news flashes & my music!
Document 15	I am spoiled by one downstairs...so got this one for upstairs to listen to music , etc. while I walked on the treadmill.

Stemming & Stopping														
Document 1	Echo	sound	same	emoji	speaker	from	Five Belo	me	sibling	has	even	one	has	Bluetoot
Document 2	Echo	respond	when	not	talk	Echo	unplug	Echo	Echo	feel	like	Echo	spy	
Document 3	not	like	almost	everytime	me	ask	Echo	question	Echo	would	talk	Echo	not	know
Document 4	you	think	about	really	not	do	much	play	music	answer	stupid	question	finally	Echo
Document 5	not	all	happy	speaker	not	great	reason	lot	time	when	play	music	Echo	will
Document 6	sad	joke	worthless											
Document 7	start	talk	Echo	self	has	not	use	Echo	since	unplug	Echo	day	not	one
Document 8	buy	Echo	healing r	play	continuou	spa	music	although	Echo	know	type	music	need	Echo
Document 9	love	Echo	dislike	can	not	sync	Apple Mus	Echo	play					
Document 10	primarily	music	bedroom	wish	speaker	better	fidelity							
Document 11	Echo	not	bad	has	issue	Echo	actually	play	want	get	Echo	use	iHeartRad	properly
Document 12	child	love	Echo	skill	amazing	like	see	Echo	get	little	smarter	answer	basic	question
Document 13	love	new	Echo	now	mainly	use	Echo	play	music	night	Echo	get	job	done
Document 14	has	three	Echo	now	even	one	barn	great	idea	when	work	horse	still	get
Document 15	spoil	Echo	downstair	get	Echo	upstairs	listen	music	etc	when	walk	treadmill		

Binary Coding	Echo	not	music	has	play	speaker	when	like	child	get	sound	one	about	talk
Document 1	1	1	0	1	0	1	0	0	0	0	1	1	1	
Document 2	1	1	0	0	0	0	1	1	0	0	0	0	0	
Document 3	1	1	0	1	0	0	0	1	0	0	0	0	0	
Document 4	1	1	1	0	1	0	0	1	0	0	0	0	1	
Document 5	1	1	1	1	1	0	1	0	1	0	0	0	0	
Document 6	0	0	0	0	0	0	0	0	0	0	0	0	0	
Document 7	1	1	0	1	0	0	0	0	1	0	0	1	1	
Document 8	1	1	1	0	1	1	0	0	0	0	1	0	0	
Document 9	1	1	0	0	0	0	0	0	0	0	0	0	0	
Document 10	0	0	1	0	0	1	0	0	0	0	0	0	0	
Document 11	1	1	0	1	1	0	0	0	0	0	0	0	0	
Document 12	1	0	0	0	0	0	0	0	1	0	0	0	0	
Document 13	1	0	1	0	1	0	0	0	0	1	0	0	0	
Document 14	1	0	1	1	0	0	1	0	0	1	0	1	0	
Document 15	1	0	1	0	0	0	0	0	0	1	0	0	0	
Total Documents	13	9	7	6	5	3	3	3	3	3	2	3	3	

Frequency Coding	Echo	not	music	has	play	speaker	when	like	child	get	sound	one	about	talk
Document 1	3	1	0	2	0	1	0	0	0	0	1	1	1	
Document 2	5	1	0	0	0	0	1	1	0	0	0	0	0	
Document 3	4	3	0	1	0	0	0	1	0	0	0	0	0	
Document 4	1	1	1	0	1	0	0	1	0	0	0	0	1	
Document 5	1	2	1	1	1	0	1	0	1	0	0	0	0	
Document 6	0	0	0	0	0	0	0	0	0	0	0	0	0	
Document 7	4	2	0	1	0	0	0	0	2	0	0	1	1	
Document 8	4	2	2	0	1	1	0	0	0	0	2	0	0	
Document 9	2	1	0	0	0	0	0	0	0	0	0	0	0	
Document 10	0	0	1	0	0	1	0	0	0	0	0	0	0	
Document 11	4	2	0	1	2	0	0	0	0	0	0	0	0	
Document 12	2	0	0	0	0	0	0	0	1	0	0	0	0	
Document 13	3	0	1	0	1	0	0	0	0	1	0	0	0	
Document 14	1	0	1	1	0	0	1	0	0	1	0	1	0	
Document 15	2	0	1	0	0	0	0	0	0	1	0	0	0	
Total Occurences	36	15	8	7	6	3	3	3	4	3	3	3	3	

tf-idf Coding	Echo	not	music	has	play	speaker	when	like	child	get	sound	one	about	talk
Document 1	0.008	0.009	0.000	0.033	0.000	0.029	0.000	0.000	0.000	0.000	0.036	0.029	0.029	0.00
Document 2	0.024	0.017	0.000	0.000	0.000	0.000	0.054	0.054	0.000	0.000	0.000	0.000	0.000	0.05
Document 3	0.015	0.039	0.000	0.023	0.000	0.000	0.000	0.041	0.000	0.000	0.000	0.000	0.000	0.04
Document 4	0.003	0.009	0.014	0.000	0.020	0.000	0.000	0.029	0.000	0.000	0.000	0.000	0.029	0.00
Document 5	0.003	0.019	0.014	0.017	0.021	0.000	0.030	0.000	0.030	0.000	0.000	0.000	0.000	0.00
Document 6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.00
Document 7	0.008	0.013	0.000	0.012	0.000	0.000	0.000	0.000	0.042	0.000	0.000	0.021	0.021	0.04
Document 8	0.010	0.017	0.025	0.000	0.018	0.027	0.000	0.000	0.000	0.000	0.067	0.000	0.000	0.00
Document 9	0.014	0.025	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.00
Document 10	0.000	0.000	0.047	0.000	0.000	0.100	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.00
Document 11	0.011	0.020	0.000	0.018	0.043	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.00
Document 12	0.011	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.064	0.000	0.000	0.000	0.000	0.00
Document 13	0.014	0.000	0.025	0.000	0.037	0.000	0.000	0.000	0.000	0.054	0.000	0.000	0.000	0.00
Document 14	0.004	0.000	0.019	0.023	0.000	0.000	0.041	0.000	0.000	0.041	0.000	0.041	0.000	0.00
Document 15	0.011	0.000	0.030	0.000	0.000	0.000	0.000	0.000	0.000	0.064	0.000	0.000	0.000	0.00

Stemming & Stopping and Coding