

Understanding the influence of the day of the week in the reviews written using SAS® Enterprise Miner™ and SAS® Sentiment Analysis Studio

Sujal Reddy Alugubelli, Oklahoma State University

ABSTRACT

Understanding customers' needs is the most critical aspect in every business. Today most companies utilize customer feedback to understand customers' desires. Analyzing what factors affect customers' feedback is therefore important for any company. This paper analyses the effect of day of the week in the way a user reacts. SAS® Enterprise Miner™ and SAS® Sentiment Analysis Studio is used to analyze reviews written on weekdays and weekends. By understanding the differences in the opinions expressed on weekdays versus weekends, a company would have better understanding of their customer reviews. In this paper reviews written on Amazon.com for two products from Electronics product category are explored.

INTRODUCTION

In the recent times most manufacturers give utmost importance to the reviews written by users on social networking sites and ecommerce websites. It is very crucial for a firm to know the key feedback and thereby providing synthesized services to their customers ^[1]. With the huge number of reviews and comments available it is an arduous job to narrow down to the most significant reviews for the betterment of the business. One most important factor that a manufacturer could keep in mind while eliminating non-useful reviews is the effect of day of the week on the way a user writes a review.

According to a research ^[4], human beings react differently in different times. They tend to over react when they have some pressures and are more accepting in nature when they are relaxed and calm. As most individuals are stressed on weekdays because of the work load and the things that happen in a day they are bit more arrogant whereas they are more peaceful and calm on weekends. Going by this concept, in this paper analysis is done on weekday and weekend review so as to understand the contributing factors of a user comment. This can greatly aid for a company to understand if a review on a particular specification is written out of a real reason or was it just because of a user's bad mood on that day.

This way analyzing the user reviews and thereby providing synthesized services would greatly aid a company in multiple ways. First valuable point is that a company gets to know the true voice of the customers. The company would also save lot of money and effort that would have been spent on non-significant requirements. This way a company can not only attain excellent customer satisfaction but also be efficient in providing the services.

METHODOLOGY

The process that is followed in this analysis is broadly classified into following steps:

- Data Extraction using Python.
- Data Preparation and initial analysis using Base SAS.
- Initial Text analysis and filtration of the text.
- Review analysis using SAS Enterprise Miner.
- User sentiment analysis using SAS Sentiment Studio.



Figure 1. Process flow diagram

DATA EXTRACTION USING PYTHON

For the purpose of the paper, the first step is to collect user ratings and reviews written. Python web scraper is used to extract the reviews from Amazon.com [5] website and are parsed. Reviews written for two products pertaining to Cell phones accessories product category are extracted. The data has the information regarding the rating given, review written, date on which review was written and few other fields.

DATA PREPARATION AND INITIAL ANALYSIS USING SAS FOUNDATION

After extracting the data from the website, the second step is to format the data in a proper way suitable for the analysis. The day of the week needed to be obtained from the posted date and also separation of the reviews written on weekdays and weekends are to be sorted and should be classified as either positive review or a negative review.

Reviews that had a rating of 5 and 4 are classified as positive reviews and reviews with a rating of 1 and 2 are classified as negative reviews. As an initial step, ANOVA is performed so as to see the differences in the ratings given on weekdays and weekends. The results are visualized as shown below. We can clearly see from the below plots that the percentage of lower ratings (1 and 2) on a weekday is more on a weekday compared to the percentage on a weekend. This clearly shows numerically that users tend to given comparatively lower ratings on a weekday than on a weekend.

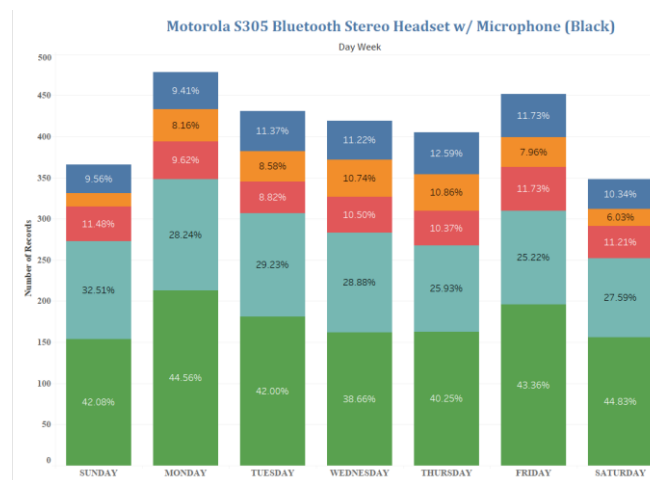


Figure 2. Product -1 Percentage of ratings on different weekdays.

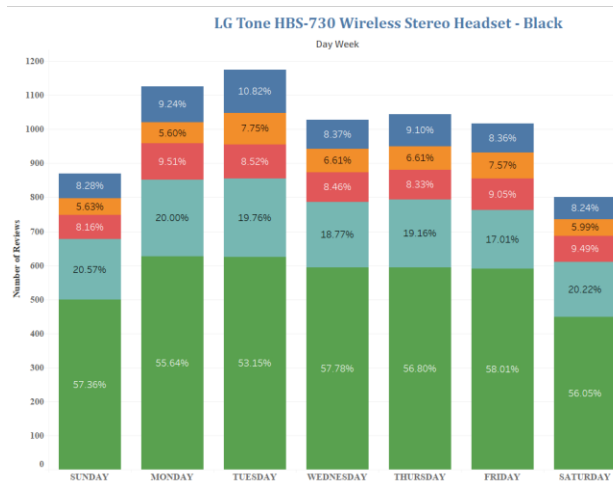


Figure 3. Product -2 Percentage of ratings on different weekdays.

TEXT MINING AND REVIEW ANALYSIS USING SAS ENTERPRISE MINER

After preparing the data suitably for text mining, data is brought into SAS Enterprise Miner. In this step initially basic text analysis is performed using text filter and text parsing nodes [3] and the frequencies of terms and documents are observed. The initial results are as follows

Term	Role	Attribute	Freq	# Docs	Keep
+ be	Verb	Alpha	2225	642N	
+ have	Verb	Alpha	1329	488N	
not	Adv	Alpha	1249	478N	
+ use	Verb	Alpha	676	400N	
+ do	Verb	Alpha	743	352N	
+ headphone	Noun	Alpha	741	327Y	
+ phone	Noun	Alpha	567	325Y	
+ work	Verb	Alpha	454	311Y	
+ good	Adj	Alpha	423	309Y	
+ quality	Noun	Alpha	394	290Y	
very	Adv	Alpha	444	287N	
+ headset	Noun	Alpha	628	285Y	
+ net	Verb	Alpha	443	273N	
+ sound	Noun	Alpha	350	272Y	
music	Noun	Alpha	418	266Y	
+ ear	Noun	Alpha	414	259Y	
sound	Adj	Alpha	269	233Y	
comfortable	Adj	Alpha	232	207Y	
+ buy	Verb	Alpha	251	203Y	
no	Adv	Alpha	270	203N	
+ great	Adj	Alpha	246	195Y	
s	Noun	Alpha	355	196N	
+ price	Noun	Alpha	220	183Y	
+ listen	Verb	Alpha	228	181Y	
bluetooth	Noun	Alpha	253	173Y	
+ call	Noun	Alpha	243	172N	
great	Adv	Alpha	186	170Y	
also	Adv	Alpha	236	169N	
+ wear	Verb	Alpha	209	160Y	
+ go	Verb	Alpha	193	155N	
battery	Noun	Alpha	162	149Y	
sound quality	Noun Group	Alpha	170	148Y	
so	Adv	Alpha	181	147N	
good	Noun	Alpha	173	146Y	
+ recommend	Verb	Alpha	154	146Y	
just	Adv	Alpha	183	139N	
+ love	Verb	Alpha	157	138Y	
+ hour	Noun	Alpha	178	137Y	
+ make	Verb	Alpha	183	137N	
life	Noun	Alpha	142	136Y	

Figure 4. Results of Terms using Text parsing node SAS EM

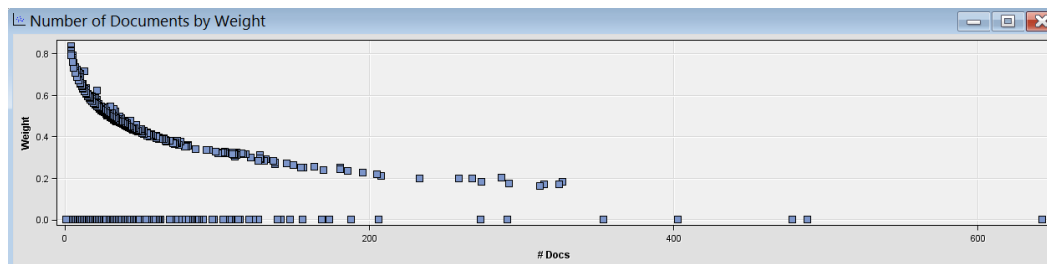


Figure 5. Results of Number of Documents by weight using Text filtering node SAS EM

REVIEW ANALYSIS USING SAS ENTERPRISE MINER

Once the initial filtration and analysis is done the following steps are followed so as to analyze the reviews written and extract the information that would be useful for eliminating the non-significant reviews and also determine the effect of day of the week on the reviews written. This step is the most crucial and the core process of the analysis.

- As the reviews are already classified as positive and negative within weekday and weekend, the negatives reviews written on a weekday are of utmost concern. This dataset is analyzed using the text cluster, text topic and concept links.
- By performing the above mentioned process few significant and moderately important topics are picked.
- Now the weekend positive reviews are analyzed using the text cluster, text topic and concept links. The previous picked topics from weekday negative reviews are looked for presence in this weekend positive reviews.

- After comparing, if the significant negative comments present in the weekday reviews are also present in the positive reviews of the weekend then most likely those reviews are because of the mood of the customer.

By following the above steps, it yields the results and helps us better understand the true voice of the customers. In addition to the analysis performed we could also perform sentimental analysis using SAS Sentiment Analysis Studio. This can clearly depict what a customer exactly feel on a weekday and what does he have in mind on a weekend.

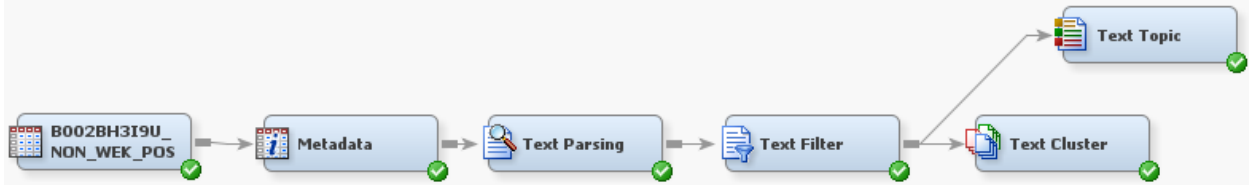


Figure 6. SAS Enterprise Miner model for the analysis

USER SENTIMENT ANALYSIS USING SAS SENTIMENT ANALYSIS STUDIO

In order to see what a user in general believes on a weekday and if it is different from weekend, sentiment analysis is performed using SAS Sentiment Analysis. By this one can get fair understanding if the reviews written are different on both the cases or are the same.

Using the current data the results of sentiment analysis is as below.

■ Positive ■ Negative ■ Overall

BEST MODEL is Smoothed Relative Frequency and Chi Square



Figure 7. Best model selected by the sentimental analysis studio

	Positive	Negative	Neutral	
	Type	▽	Body	Weight
1	CLASSIFIER	great		0.046423
2	CLASSIFIER	price		0.039659
3	CLASSIFIER	music		0.024507
4	CLASSIFIER	sound		0.024349
5	CLASSIFIER	phone		0.023998
6	CLASSIFIER	comfortable		0.023585
7	CLASSIFIER	love		0.021847
8	CLASSIFIER	call		0.018949
9	CLASSIFIER	listen		0.016861
10	CLASSIFIER	easy		0.015936
11	CLASSIFIER	clear		0.013647
12	CLASSIFIER	Bluetooth		0.013484
13	CLASSIFIER	best		0.013289
14	CLASSIFIER	always		0.013184
15	CLASSIFIER	far		0.013174
16	CLASSIFIER	Great		0.011991
17	CLASSIFIER	cell		0.011923
18	CLASSIFIER	happy		0.011455
19	CLASSIFIER	life		0.011339
20	CLASSIFIER	wife		0.010981
21	CLASSIFIER	value		0.010981

Figure 8. Statistical rules formed to classify a review as positive

	Positive	Negative	Neutral	
	Type	▽	Body	Weight
1	CLASSIFIER	stop		0.067178
2	CLASSIFIER	return		0.064698
3	CLASSIFIER	poor		0.039243
4	CLASSIFIER	disappoint		0.029197
5	CLASSIFIER	week		0.027773
6	CLASSIFIER	send		0.025152
7	CLASSIFIER	waste		0.024742
8	CLASSIFIER	not		0.023669
9	CLASSIFIER	warranty		0.023093
10	CLASSIFIER	Then		0.022898
11	CLASSIFIER	minute		0.022034
12	CLASSIFIER	month		0.020863
13	CLASSIFIER	red		0.020676
14	CLASSIFIER	working		0.017632
15	CLASSIFIER	completely		0.017486
16	CLASSIFIER	stopped working		0.017265
17	CLASSIFIER	replacement		0.016953
18	CLASSIFIER	useless		0.01647
19	CLASSIFIER	plan		0.016339
20	CLASSIFIER	refund		0.014933

Figure 9. Statistical rules formed to classify a review as positive

RESULTS

- In the current scenario, for the product Motorola Stereo Headset, the term volume seems to be one of the significant topic in negative weekday reviews. The same term appears in the weekend significant positive reviews list too.
- This clearly raised the flag of uncertainty of the proper feedback. For this purpose, the reviews which consisted of that term are pulled and are analyzed.
- After analyzing the reviews, it is clearly seen that there is actually no problem with the volume aspect of the device. So the negative reviews written on weekday regarding volume are only because of bad mood of the customer and not really a problem of the product.
- Similarly, with the second product LG Tone HBS-730 wireless stereo headset. There were negative reviews written about earbuds on weekdays but lot of positive feedback is written on weekend. Thus it clearly shows the presence of day of the week effect on the way people write a review.

So, it seems that there is an effect of the day of the week on the way people write a review. By performing such analysis, companies can better understand the nature of the reviews.

MODEL: LG TONE HBS-730 WIRELESS STEREO HEADSET – BLACK

Cluster ID	Descriptive Terms	Frequency
1	+volume +loud +low set' +volume control' +low +control +ios +set +iphone +music +comfortable +listen +device +button +read	45
2	+pocket +foot +phone +signal +walk +drop +cut +bluetooth +connection +range +lose +listen +pair +keep +headphone	104
3	audio +support +issue +hbs-700 +connection +music +iphone +battery +ios +device +connect +comfortable different +second +know	116
4	+charge +charger +full +minute +replacement first +day +unit +turn +keep +hold right +work +package amazon	63
5	+model 'previous model' previous +version +old +battery +love +range +second +sound quality' +year +hbs-700 back +cut +issue	33
6	+box +package real +fake +open amazon +order +original +charger +receive +cheap +look +item different +warranty	75
7	+hear +noise people +talk money +problem +work well +item +want +stop +know +button +set back	115
8	+month +bud +ear bud' +ear +warranty +break +wire +stop +last +year +piece +long +design +cheap +work	90
9	+quality +sound quality' +sound +range +horrible +bluetooth +hope +bluetooth headset' +review +read +device +good lg +headset +tone	62

Figure 10. Clusters formed for negative reviews on Weekday

Cluster ID	Descriptive Terms	Frequency
1	+headphone +wire +review +issue +range +problem +set +want +pair +little +fit +neck +order +buy well	476
2	+noise +volume +voice +hear +connect +device well +little +thing 'sound quality' +range +time +excellent +quality +problem	423
3	+battery +life +battery life' +good +long +quality +sound +great +comfortable 'sound quality' +amaze +range +easy +excellent +far	176
4	+phone +music +love +listen +hand +drive +talk +free cell +product able +work +great +great product' +clear	221
5	+battery +day +last +charge +hour 'long time' +long +week +work +time +listen +wear +comfortable +great product' +music	148
6	+love +great +recommend +product +buy +purchase +husband +friend +gift family +order +price +great product' +item +awesome	339
7	lg +tone +lg tone' +headset +bluetooth +look +review cell +recommend +friend +long +talk +highly +item +voice	92
8	+ear +bud +ear bud' +piece +stay +neck +fit +little +hold +wear +happy +item +comfortable +thing +hear	177

Figure 11. Clusters formed for positive reviews on Weekend

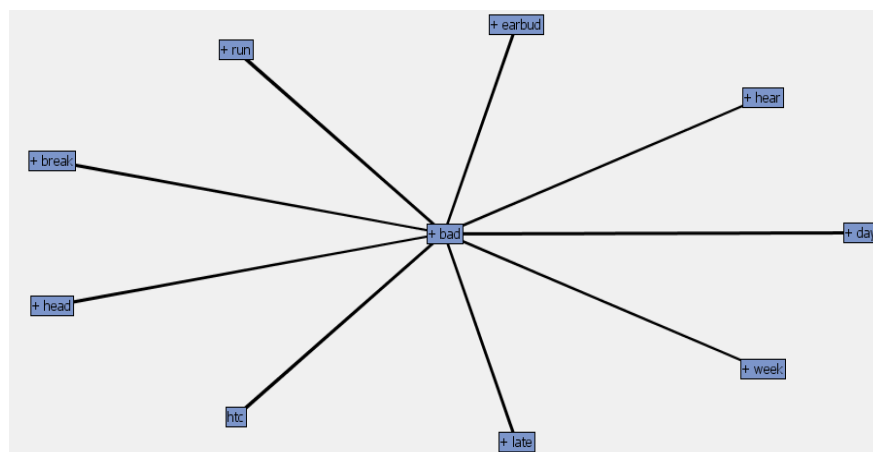


Figure 12. Terms connected to “Bad” on weekday

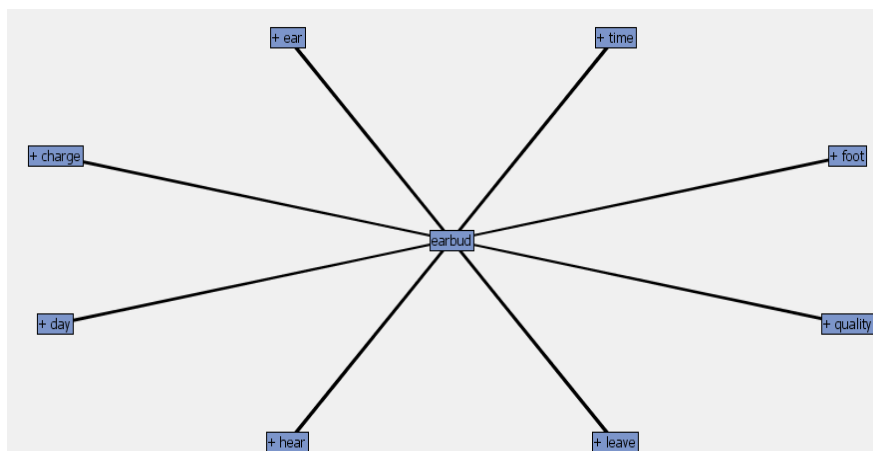


Figure 13. Terms connected to “Earbud” on weekend

I LOVE these headphones. The first pair I received was defective - the battery got so hot it burned my neck and it would die after 40 minutes of play. I ordered a replacement and it works beautifully. They are so comfortable I forget I am wearing them - which is exactly how I want wireless headphones to feel. The earbuds are comfortable and deliver great sound quality. The controls are easy to memorize and respond quickly. I am thrilled with this purchase.	5	They play music from various bluetooth devices. They get fairly loud, block exterior noises and the battery lasts for the entire day. Keep them away from the cat. The dangling earbuds are really tempting to the furball.	5
---	---	--	---

Figure 14. Examples of positive reviews about earbud on Weekend

because of the tiny wiring to the earbuds, the wire breaks internally and had a friend cut the wires and resolder them. it happened again a few months later. nice idea but has a design flaw.	1	Was satisfied with my headset until this morning I look up when my music kept going in and out to see the left side ear bud cord has broken near the connection of the headset. The earbud cords are stiff and this could have come from placing the bud in my ear. Not happy at all. I haven't even had this item 6 months.	2
---	---	---	---

Figure 15. Examples of negative reviews about earbud on Weekday

MODEL: MOTOROLA S305 BLUETOOTH STEREO HEADSET W/ MICROPHONE (BLACK)

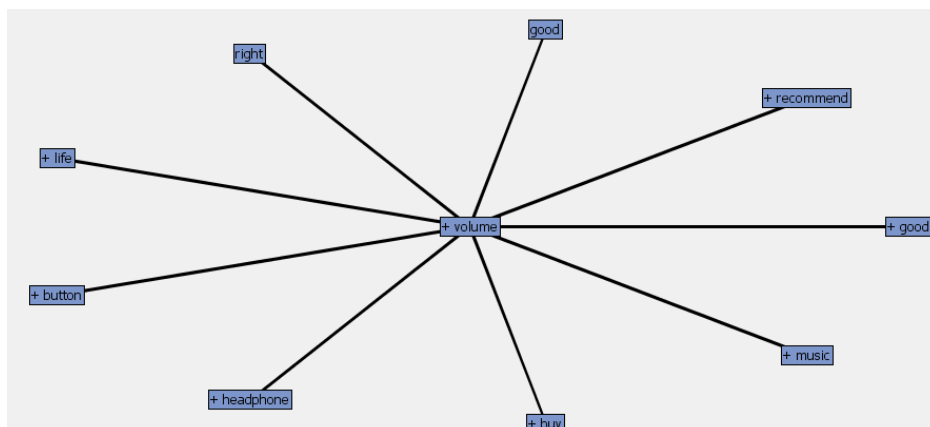


Figure 16. Terms connected to Volume on Weekend

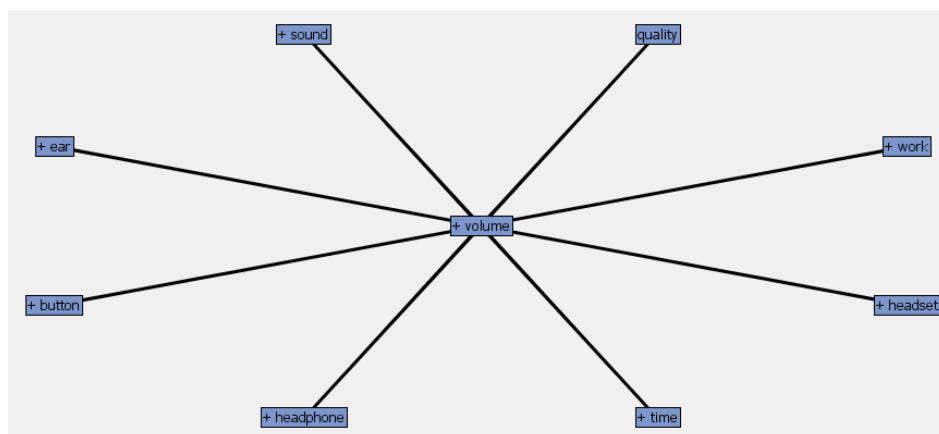


Figure 17. Terms connected to Volume on Weekday

Cluster ID	Descriptive Terms	Frequency
1	listen +music +'battery life' able +pair +battery +purchase back first amazon +motorola +life +device +problem +feel	37
2	device bluetooth price +recommend +issue +keep +headset +phone +great motorola quality +bluetooth +bad +find +connect	56
3	arrive +mac +iphone +pretty amazon sound +connect +'sound quality' +day quality +easy return +stop +connection +feel	29
4	loud +hear +ipod +noise +laptop +finally +sound +little +phone +comfortable +connect +fit +connection +bluetooth +return	32
5	volume +button +'volume button' +hit +break +control +low +look +useless +replace +easy +month +drop +hard +long	28
6	item +waste money +fail return +window amazon +purchase +week +replacement +time +device +connection +second working	31
7	red 'turn on' +charge +last +running +stop +sweat +month working +light +workout great gym few +week	92
8	wear +head +uncomfortable +ear +minute +fit +hour +hard +comfortable +hear +phone +design +music +noise +feel	52

Figure 18. Clusters formed for negative reviews on Weekday

Cluster ID	Descriptive Terms	Frequency
1	touch +ipod +'ipod touch' ipod +pause +play +audio +button +volume +control +wire back +turn well +computer	80
2	hear +headphone +'bluetooth headphone' +price +first +work +pair great +set +music +bluetooth +little +great +sound +listen	104
3	headset +'bluetooth headset' +listen +music +look +excellent +want +hear +bluetooth +s305 bluetooth +phone +purchase +player +recommend	94
4	battery +life +'battery life' +'sound quality' sound +quality +easy +excellent +last +set +hour +iphone +charge +love +buy	57
5	battery +charge +device +day +hour +last +life +iphone +'bluetooth headphone' +connect +'battery life' +pair +headphone +problem +button	201
6	motorola +bud +s305 +ear +wear +head +find +comfortable +fit +recommend bluetooth +audio +listen +control +set	81
7	product +great +price +love +sound +gym +recommend great +fit +good +purchase +thing +buy +excellent +range	134
8	tooth +blue +'mp3 player' +player +answer running able +exercise +wire +phone +problem +audio right +computer +pause	39
9	watch +movie +tv +ipad +computer bluetooth +play +turn +motorola +audio +range +device +player +find +want	45

Figure 19. Clusters formed for positive reviews on Weekend

Price was good, but these hurt my ears, and the audio had very high treble, so when I turned up the volume it hurt my ear drums. I was getting these to replace my RocketFish RF-MAB2, but it turns out that the Rocketfish is much better. They have a graphic equalizer built in and they fit much more comfortably. I'm sticking with those.	2	Okay so this thing works. It works fine as a bluetooth device and a headphone. Having said that, there are several issues. The quality is just not upto par. I'm not expecting audiophile quality but c'mon, it's noisy and it at times picks up other signals and noise. The volume is also anemic and thin. The whole thing feels a tad shabby in terms of construction. I returned these immediately.	2
---	---	--	---

Figure 20. Examples of negative reviews about volume on Weekend

SOUND IS GREAT THE VOLUME DOESNT GO THAT LOUD BUT GREAT SURROUND AND BASS ON MY MUSIC WHICH IS JAZZ AND OLD SCHOOL AND ONE AWESOME THING IS THE FLEXIBILITY OF THE HEADPHONE I HAVE DROP THEM ABOUT 10 TIMES SMASH THEM IN MY BOOKBAG AGAINST MY BOOKS BY MISTAKE OF COURSE EVERYDAY AND THEY STILL WORK AND LOOK BRAND NEW.	5	So far I love the quality of sound both for calls, and for listening to music, or audio books. The volume control is great because it emits a tone change so that you know if you are increasing or decreasing the sound. Over-all good quality at a great price!	4
--	---	---	---

Figure 21. Examples of positive reviews about volume on Weekend

CONCLUSION

This paper illustrated the influence of day of the week on the way a customer writes a review. Specifically, in the two examples, data is extracted and prepared for text analytics, performed text analytics, prepared data for exploration and reporting. Manufacturers can leverage this analysis or elements thereof, in their own analysis.

In this paper I explored the behavior changes observed using text analysis. Text cluster node and text topic node helped in finding out the significant review written by the customer. Text filter node helped in finding out the terms that are related to those significant topics. By this way a company can greatly synthesize its services and optimize the efforts spent on the betterment of the product.

FUTURE SCOPE AND LIMITATIONS

- The analysis performed in this paper does not take into consideration of the nature of the customer. So if the type of a customer could be known then evaluation of the results would make much more sense.
- The paper applies to a working customer who has workload on Monday through Friday and has an off from work on Saturday and Sunday.
- May be product from product categories specific to certain age groups could be analyzed so as to understand if there is an age factor. For example, mostly video games products are used by younger generations and home appliances are purchased by elder people so products from those products categories would help us understand if there is any age factor on how the behavior changes according to day of week.
- The length of the review written also has a significance in the intension of the review written. Survey tells us that when customers write just a short but harsh comment then in most cases it is not a true one and biased by some factor. So we can add the length of a comment factor so as to tune the comments.
- Time of the day is one more important factor which need to track of. Comments would be better in the early part of the day than in later part by following the concept of change in behavior because of work pressures. So if we could bring in the time factor that would greatly aid the company

REFERENCES

1. Thomas Y. Lee, "Needs-Based Analysis of Online Customer Reviews".
2. Tom Sabo, SAS Federal LLC "Uncovering Trends in Research Using Text Analytics with examples from Nanotechnology and Aerospace Engineering".
3. Chakraborty, Goutam and Pagolu, Murali Krishna "Analysis of Unstructured Data: Applications of Text Analytics and Sentiment Mining".
4. Episode 36: Science of Deception on June 28 2016 from <http://www.npr.org/podcasts/510308/hidden-brain/>.
5. Amazon product data, Julian McAuley <http://jmcauley.ucsd.edu/data/amazon/links.html>.

ACKNOWLEDGMENTS

I wish to express my deep sense of gratitude and indebtedness to the Director of Master of Science in Business Analytics, Dr. Gotham Chakraborty, Spears School of Business, Oklahoma State University for the good lectures on analytics and for his inspiring guidance, constructive criticism and valuable suggestion throughout my project. Without his support I would have not able to be so successful during my internship.

I would also like to express my gratitude to Dr. Miriam McGaugh for her valuable suggestions and encouragement at various stages of my project.

RECOMMENDED READING

- *Chakraborty, Goutam, Pagolu, Murali Krishna and Satish Garla* Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS
- https://www.sas.com/en_us/whitepapers/evolution-of-analytics-108240.html

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Sujal Reddy Alugubelli
Oklahoma state University
(405)385-2039
sujalreddyalugubelli@gmail.com, sujal_reddy.alugubelli@okstate.edu
<https://www.linkedin.com/in/sujal-reddy-alugubelli-08651093>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.