

April 19th Presentation

Sentiment & Topic Modeling

Session Agenda

1

What are these concepts?

2

How Sentiment Analysis & Topic Modeling Work

3

Research Approach & Caveats

4

Verification & Validation Procedures

5

Visualizing the Resulting Data Elements

Introductions



Larry Mallak, Ph.D.

Professor of Industrial and Entrepreneurial Engineering & Engineering Management
Director of Engineering Management Research Laboratory

Applied academic researcher integrating social science methods into engineering, technical, and healthcare domains. Primary research areas of organizational culture, workplace resilience, and ethnographic practices for new product development.



Nolen Akerman, Ph.D.

Data Scientist, First Analytics <https://public.tableau.com/app/profile/nolen.akerman>

Experienced data scientist & researcher with over twenty years working with companies in the med-tech & consumer packaged goods sectors. Has partnered with Larry Mallak on research applying unstructured data analysis techniques for evaluation of product advantage components, organizational culture, and service quality.

What are these concepts?

What is unstructured data?



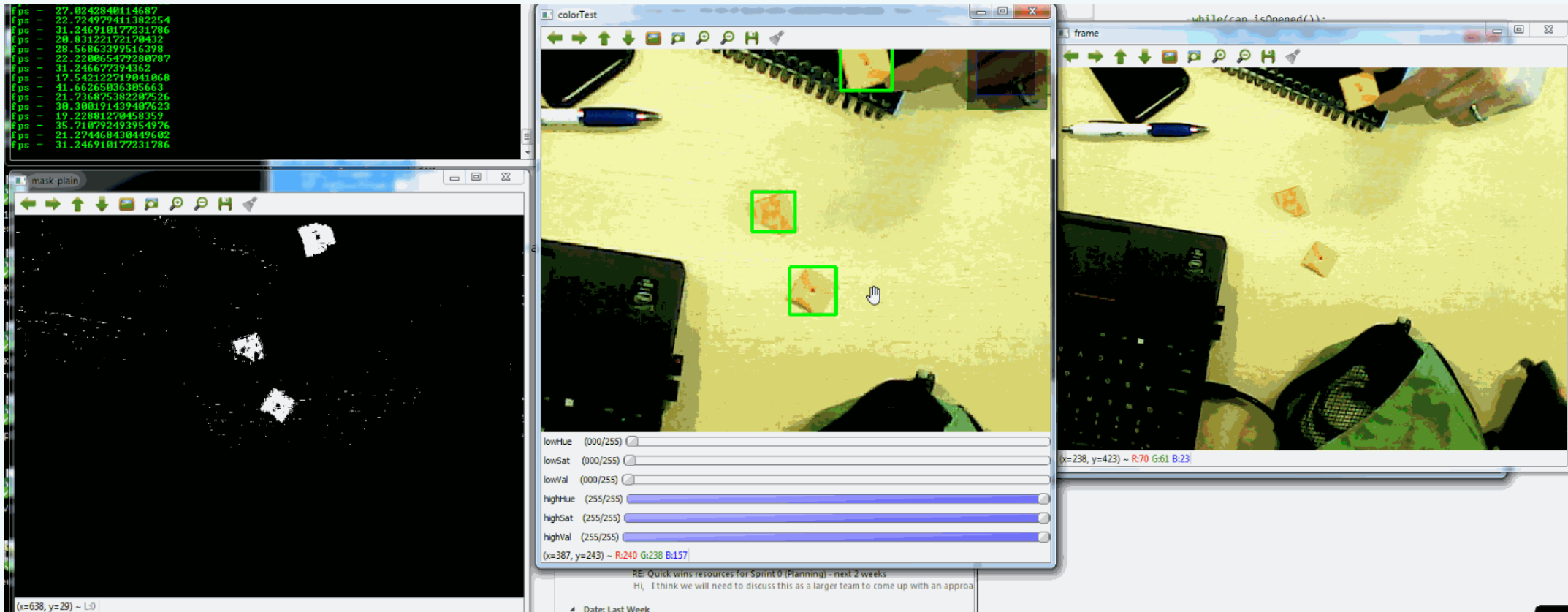
Unstructured Data

- Data is dispersed without a clear model definition
- Involves some operations to become structured
- Can contain numeric, text, audio, or video data

Examples:

- Unstructured documents (pdf, word)
- Pictures
- Videos
- Voice Recordings
- Blog Posts
- Reviews
- Surveys

What is unstructured data?



Video is one form of unstructured data

- You can structure this video data to count 'CheezIts', determine color (over cooked), count number of broken 'CheezIts', and generate batch analysis of quality when collated.
- This same methodology can be applied to text data, picture data, audio data, or a variety of unstructured sources.

Advantages of unstructured data

"Information Richness"

★★★★★ **Good stuff!**, April 18, 2012

By [~Kimber~](#) (MA, USA) - [See all my reviews](#)

VINE™ VOICE

This review is from: Kellogg's Krave Chocolate Cereal, 11.4-Ounce (Pack of 4) (Grocery)

Vine Customer Review of Free Product ([What's this?](#))

My daughter has been begging me to buy this cereal, so I was excited to get to try a box to review.

Upon pouring my first bowl of cereal, I was a little surprised.. I hadn't paid a lot of attention to the commercials, so I had pictured this more like shredded wheat with chocolate in the middle. It's actually a golden crispy cereal on the outside with a chocolate center.

About the cereal itself, it's quite tasty, the golden crisp nicely offsets the yummy chocolate blast in the middle. It does get a little soggy in milk, but it generally doesn't last long enough to be a real concern. It was a huge hit in my household, I only got one bowl, when I went back for more, it was gone! Since then, by popular demand, I purchased the double chocolate Krave. While the chocolate is too intense for me, my husband and daughter love it. I must now keep a box of Krave on hand at all times or suffer my family's wrath!

Advantages of Unstructured Data

- Provides context to a rating (5/5 stars – but why?)
- Provides additional 'meta-data' – location, date, family, associated comments
- Collated with additional comments, provides a repository of drivers of positive and negative sentiment.

What is Sentiment Analysis & Topic Modeling?

Sentiment Analysis

- Sentiment analysis is a type of unstructured data analysis
- It involves detecting the level of 'positive' or 'negative' sentiment in written text
- Main objective is to use written text as a proxy for level of positive or negative emotion associated to the experience from the respondent
- Uses a standard or domain specific library of terms and inversion factors to identify positive and negative words or phrases

Topic Modeling

- When we identify positive or negative comments, we want to know what specifically they are referring to
- Topic modeling allows us to group together similar responses based on the correlation of the words used

How do these methods work?

How sentiment analysis works

“ I appreciate the way employees will greet one another. (usually with a smile and a 'hello!') This comment includes the CEO, the doctors and our nurses. I find this encouraging! ”

“ Keeping employees on system and shifting them when they perform poorly to the next department ”

‘poorly’ = negative, strongly subjective, value = 9.03.
Resulting calculation: Positive = 1.03, Negative = 9.43

$$\text{Rating}(\text{negative}) = \left\{ \sum_1^n \text{abs} \left(\text{LN} \left((p_{\text{strong}}|p_{\text{weak}}) \times \frac{\text{Prior}}{4,175} \right) \right) \right\} + \text{abs} \left(\text{LN} \left(\frac{4,175}{6,499} \right) \right)$$
$$\text{Rating}(\text{positive}) = \left\{ \sum_1^n \text{abs} \left(\text{LN} \left((p_{\text{strong}}|p_{\text{weak}}) \times \frac{\text{Prior}}{2,324} \right) \right) \right\} + \text{abs} \left(\text{LN} \left(\frac{2,324}{6,499} \right) \right)$$

Sentiment Analysis

- A variety of different methods can be grouped generally into ‘bag of words’ approaches and supervised learning approaches.

Examples

“ The simple but striking visuals and text combined with the excellent speaker and moderators were clear, consistent, and also engaging! ”

Highly Positive 2.425 Sentiment

34.57 Positive; 1.00 Negative with 20 word count.

“ Way too many people in the chat box, it was hard to read and follow.”

Highly Negative .461 Sentiment

1.0 Positive; 22.10 Negative with 15 word count.

Has difficulty with mixed sentiment

“ I sit through a LOT of boring webinars where I may find one or two nuggets, not this one.”

Negative .595 Sentiment

1.0 Positive; 7.98 Negative with 19 word count.

How topic modeling works

Cost
lasted
lasting
lasts
Pay
\$\$
a steal
Afford
Bang for
Bang for
bargain
Best buys
Best Deal
Budget
charge you
Cheap
Cheaper
Clearance
closeout
Discount
do not last
does last
does not last
Dollars
don't need much
economical
Expense
Expensive
for less

- Generally, there are two ways to approach both Sentiment Analysis & Topic Modelling – **Classification** using rules & lists or **Clustering** using math and dynamic association
- Sample Classification List for 'Value'

Verbatim	Human Coded Sentiment	
	Sentiment	Sentiment TRI
Almost \$10 per bottle is outrageous	Dissatisfied	NEGATIVE
At \$250 this is a good buy, but at the regular price of over \$400 I would not buy it	Slightly Negative	NEGATIVE

Topic Modeling

- In research settings, you can approach Topic Modeling in two general categories:
 - A.) **Exploratory Analysis:** identifying key trends without prior hypotheses. (Critical Incident Technique initial coding). Better aligned to “Clustering” – unsupervised.
 - B.) **Confirmatory Analysis:** grounded theory approach based on preconceived constructs. Better aligned to “Classification” – supervised.

Topic Model Example Model Output

Examples of Topic Model Identification

Topic Group 4:
“Presentation Artifacts”

Topic Group 5:
“Perceived Utility of Learning”

*“Excellent materials, practical information,
clear and engaging speaker .”*

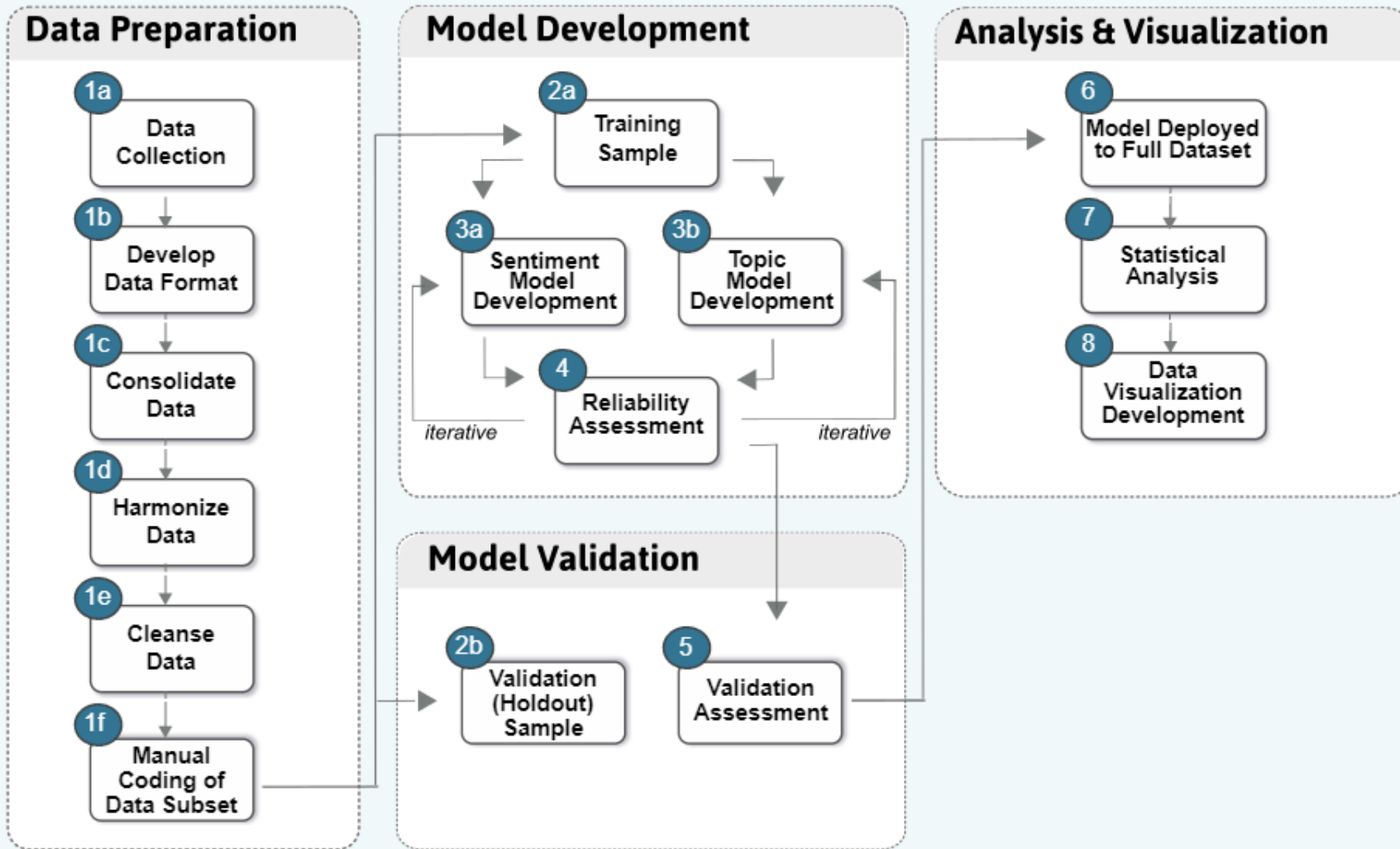
Topic Group 3:
“Presenter”

Coding Multiple Topics

- Main objective is to identify ‘themes’ (exploratory) or properly cluster (confirmatory) relevant topics
- Uses a standard or **domain-specific** text matrix is used to classify opinion phrases
- Often several different ‘topics’ are referenced in the same ‘opinion phrase’

Research Approach and Caveats

General Methodology



Unlike a 'black box' our approach provides value through:

- Curated and customized word matrix & sentiment database
- Research based topic model development (guided approach)
- Addition of traditional statistical techniques to the numerical results
- Full validation assessment for intended use of models to identify bias or overfitting
- Clear designation of the generalizability of results

Important Considerations

- These techniques benefit from the nature of large data (some classifications might be incorrect, but with a large enough dataset the signal is amplified and the noise is decreased)
- You should not rely on a 'black box' general model available on the internet for your classifications (see side panel) →
- You should provide statistical 'scaffolding' to provide face validity and rigorous validation of your models
- You still need to invest time! The benefits of the model are on the reusability. You may actually spend more time upfront developing the model and receive the payback over time.

Importance of Context

Sentiment

Negative (Food)

This frozen pizza tastes **gross**!

Neutral

I ordered a **gross** of pencils.

Topic

Quality Concern

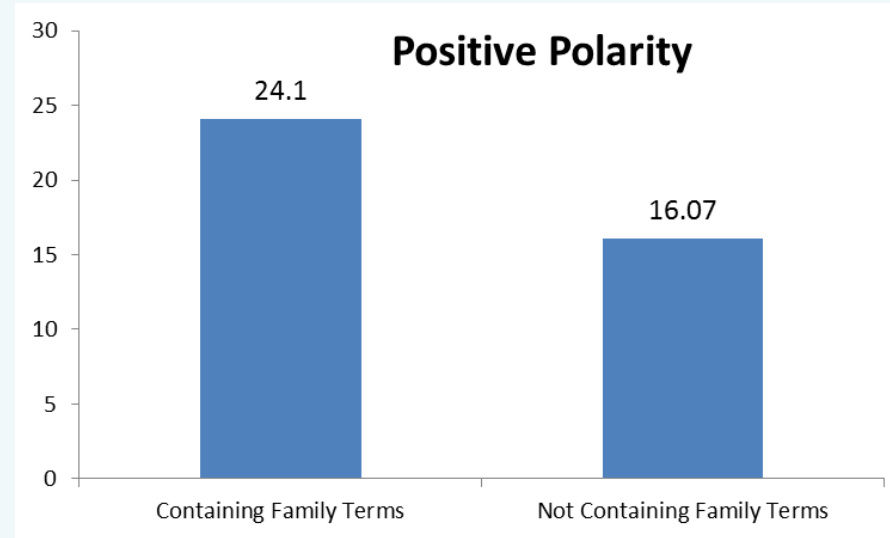
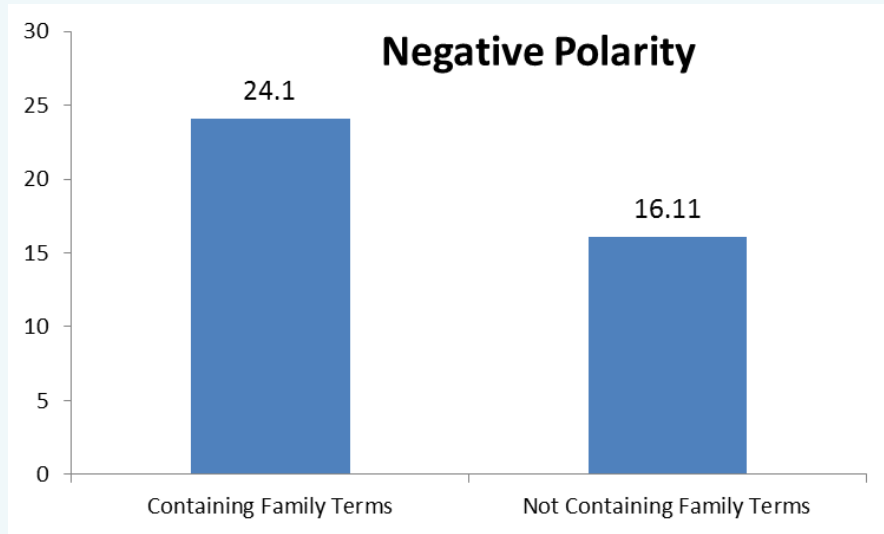
My computer emits a **loud sound**

Positive Product Feature

Speakers provide a crisp **loud sound**.

Topic Modeling Example

Case Study: That Experience was Mine: The Effects of Family Associations on Customer Perceptions



Topic Modeling to Test Hypotheses

- Do negative or positive experiences involving family members (versus single guests) influence the magnitude of negative or positive sentiment?
- What topics elicited the most positive or negative sentiment? (ANOVA)

Verification & Validation Procedures

Model Verification & Validation

Verification (training sample)

- Don't "overfit" the model
- Use weighted or normal kappa calculation for inter-rater reliability
- 5-10% coded responses from dataset
- 80%+ above accuracy a good goal
- >.70 weighted kappa a good goal

Validation (Holdout Sample)

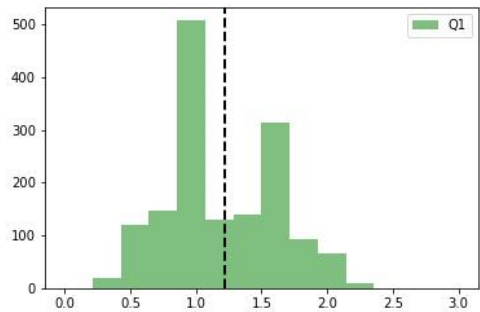
- Data 'unseen' by the model
- Attempts to minimize 'overfitting'
- 5% of coded values
- 75%+ accuracy a good goal
- >.65 weighted kappa a good goal

RESULTS					
		Model			
Human		Positive	Neutral	Negative	
	Positive	200	5	14	219
	Neutral	47	93	37	177
	Negative	7	3	175	185
		254	101	226	581

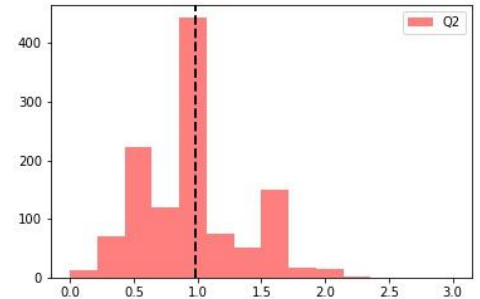
RESULTS					
		Model			
Human		Positive	Neutral	Negative	
	Positive	19	6	4	29
	Neutral	5	29	10	44
	Negative	1	2	54	57
		25	37	68	130

Fitness for Use (face validity)

It is theorized since one question is asking for positive responses and one question is asking for responses related to improvement, that the mean sentiment between the two questions would be statistically different. This is a logical and intuitive hypothesis and the results of this test are used for 'face validity' for the sentiment scoring model.



Question 1 (POS):
"What aspect of this session was especially good"



Question 2 (NEG):
"What aspect of this session needs the most improvement"

Question	n	M	SD	min	25%	50%	75%	max
Question 1	1547	1.2192	0.4268	0.2974	1.0000	1.0605	1.6432	2.4248
Question 2	1186	0.9810	0.4125	0.1522	0.6135	1.0000	1.1172	2.3023

Welch's *t*-test Results

```
from statsmodels.stats.weightstats import ttest_ind
ttest_ind(POS_Arr, NEG_Arr, usevar='unequal')
```

An independent groups Welch's *t* test, assuming unequal variances revealed Question 1 ($M=1.22$, $SD=.427$, $n_1=1,547$) differed from Question 2 ($M=.981$, $SD=.413$, $n_2=1,547$) as predicted, $t(2590.76)=14.73$, $p<.0001$

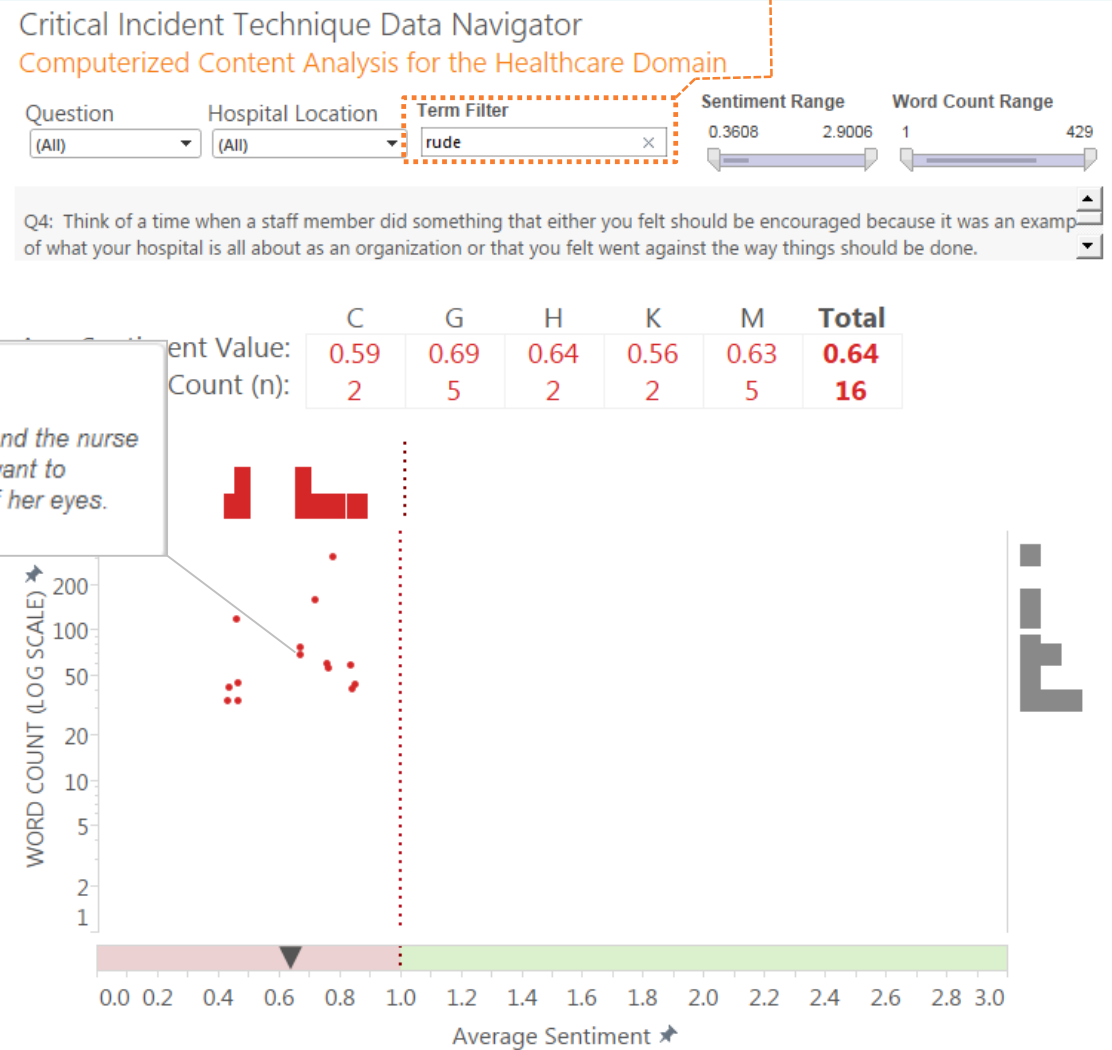
Data Visualization Demonstration

Making your Results ‘Usable’

Adding ‘structure’ to unstructured data allows for investigation of trends and identification of key topics.

“Information richness” can be found in responses with high word count and high emotional content (high or low sentiment)

A (A.) Ability to filter on key words to identify trends in drivers of satisfaction or dissatisfaction



(B.) ‘Mouse-over’ effects that allow for evaluation of the full critical incident text for individual points within the scatterplot.

Bringing it all together

Sentiment Analysis Data Navigator Computerized Sentiment Analysis of Amazon Reviews

Amazon Category: (All) Term Filter: 5 Filter for Word Count: 139 0.090 Filter for Calculated Sentiment: 3.130

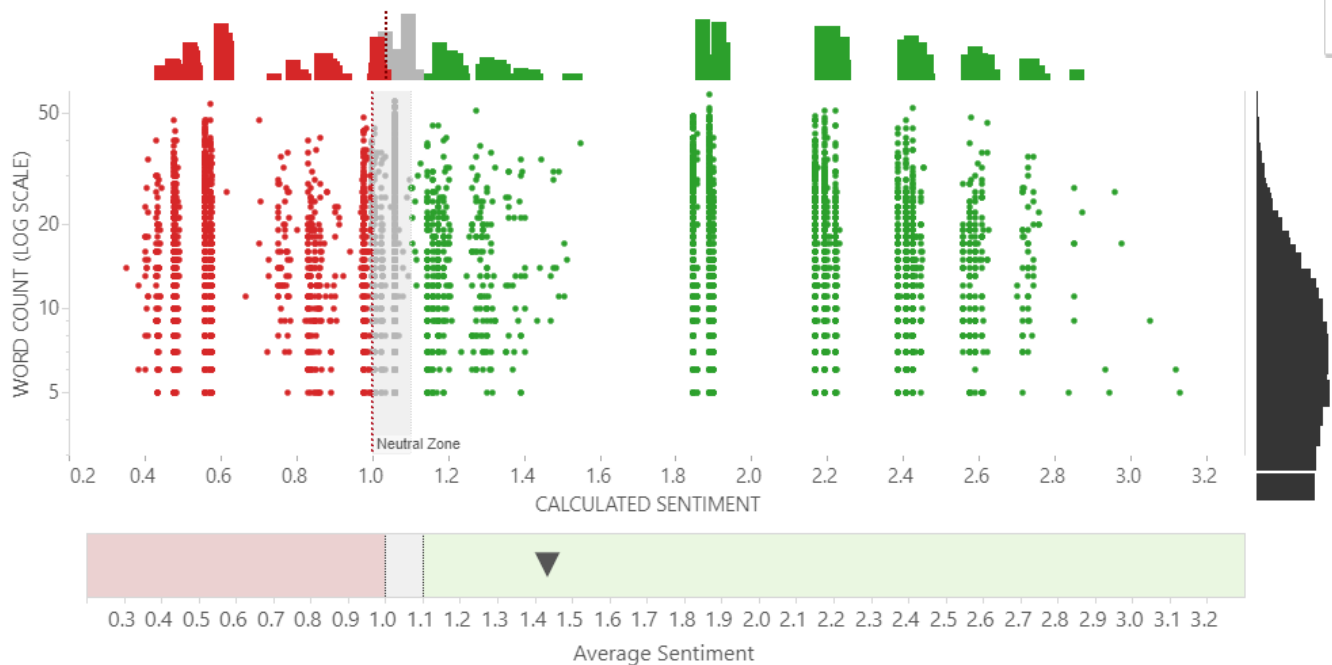
Min Threshold for Negative Sentiment: 1 Size of Buffer for Neutral Comments: 0.1

AVERAGE SENTIMENT SCORES BY TECHNOLOGY PRODUCT CATEGORIES

Camcorders	Compact Stereos	Digital SLRs	LED Monitors	Phone Docking St..	Photo Printers	Portable DVD Players	Tablets	Touchscreen Notebooks	Total
1.525	1.555	1.656	1.553	1.618	1.517	1.350	1.495	1.478	1.528

AVERAGE SENTIMENT SCORES BY COMMODITY PRODUCT CATEGORIES

Bathroom Cleansers	Chips and Crisps	Cold Cereals	Cookies	Deodorants	Glass Cleaners	Liquid Detergents	Liquid Soaps	Normal Shampoo	Total
1.570	1.729	1.657	1.698	1.540	1.685	1.581	1.677	1.512	1.627



<https://public.tableau.com/profile/nolen.akerman#!/vizhome/SentimentAnalysisNavigator/CriticalIncidentNavigator>

FILTERS

Date of Session 7/9/2009 5/27/2020 Sentiment Score 0.404 2.616 Word Count 1 82 ☒ Webinar ☒ WORKSHOP

Table 1. Opinion Phrase Counts & Sentiment by Topic Group

	Physical & Technology Characteristics	Presentation Design Characteristics	Presenter / Delivery Characteristics	Presentation Artifacts / Assets	Perceived Utility of Learning	Interaction / Interactivity Characteristics	Unclassified Opinion Phrases
Count of Phrases	276	1,010	558	782	589	913	1,834
Full Sentiment	1.02	1.28	1.42	1.36	1.45	1.32	1.19

Chart 1. Opinion Phrases Plotted by Word Count and Sentiment

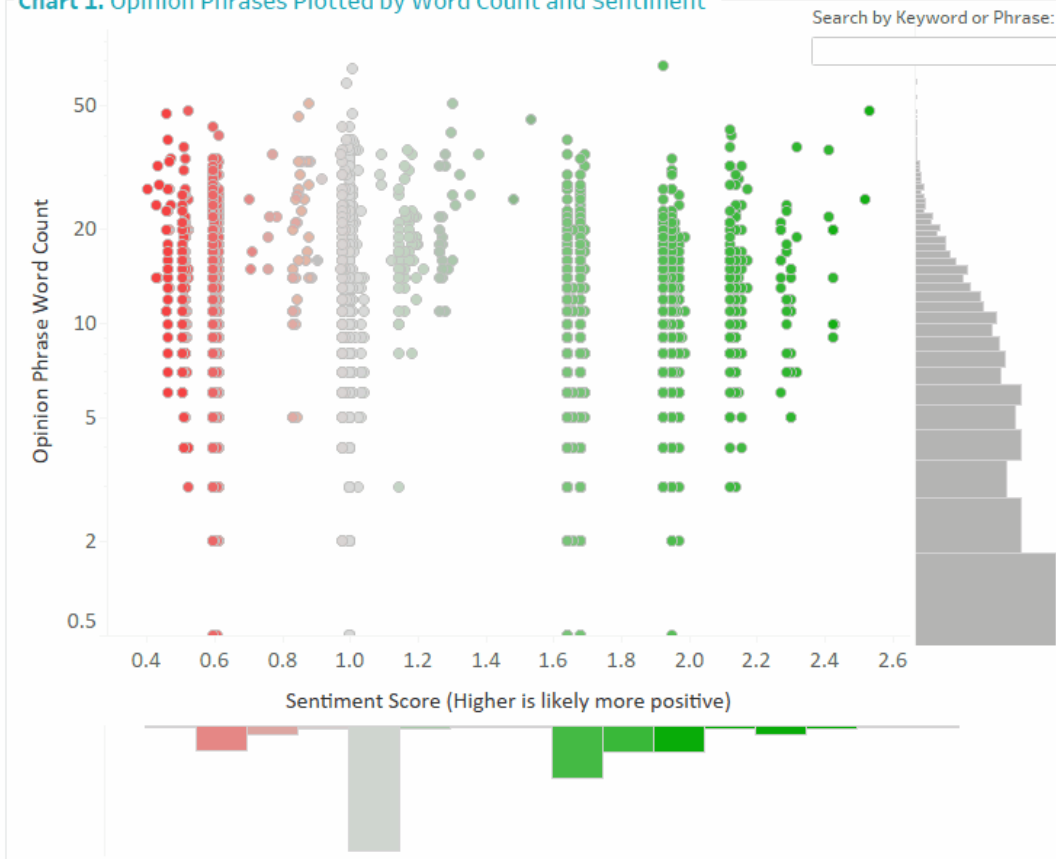


Table 2. Opinion Phrase Details

(All)	
Nice variation in speakers, well timed, paced it so that we didn't really run out of time, good visuals, facilitated so that it was engaging and easy to understand and even offered platfor..	2.531
use of interactive tool to keep us engaged, simple clear, concise presentations with good use of visual and minimal text, made it easy to follow	2.520
Simple, well organized, not too long or short, provided resources	2.429
The simple but striking visuals and text combined with the excellent speaker and moderators were clear, consistent, and also engaging	2.425
Excellent slides, excellent content, free resources, and well prepared presenters	2.425
Lori's presentation was good well organized, clear, and excellent infomration about what needs to be included in evaluation plans	2.425
The quality of the resources provided seems excellent and I will be able to easily share it with my colleagues	2.425
The webinar was organized well and the format helped keep it interesting and engaging	2.425
Well organized, good speakers, clearly informed about the subject matter	2.425
the moderator, Mike, was made some very good questions; the time was perfect; I really appreciated the fact of being linked with my google agenda, as well as the reminders; the s..	2.413
This webinar was balanced between speakers, great coverage of topics in the time allotted, terrific resources, and strong chat involvement by audience	2.413
The organization of this webinar was exceptional, even for ATE webinars the flow between speakers was perfect and the case example was genius with the lego graphics I don't know i..	2.316
Great sound, great materials, and really loved the audience participation that led to the sharing of resources amongst colleagues	2.316
Loved interactive component; excellent & engaging speakers	2.316

Example exploration identifying 'expertise' as a highly correlated term with positive views of 'presenter' characteristics

Thanks & Questions

Feel free to contact Nolen.akerman@gmail.com or mallaki@wmich.edu for any additional questions