Penn Medicine

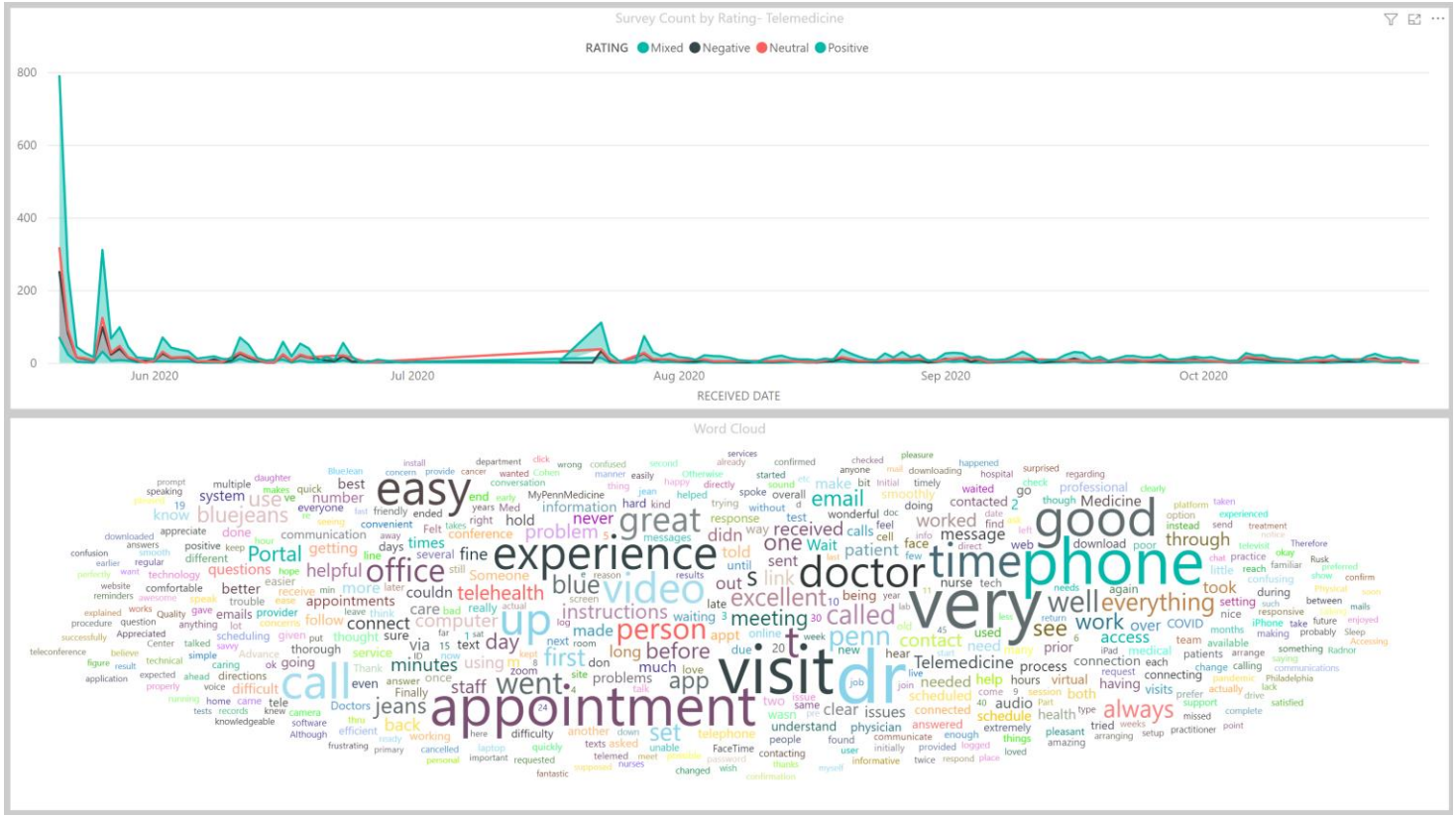
Patient Satisfaction Sentiment Analysis w/ VADER

Frank Zhao- Sr. Business Intelligence Analyst
Department of Medicine

Sept 28, 2020

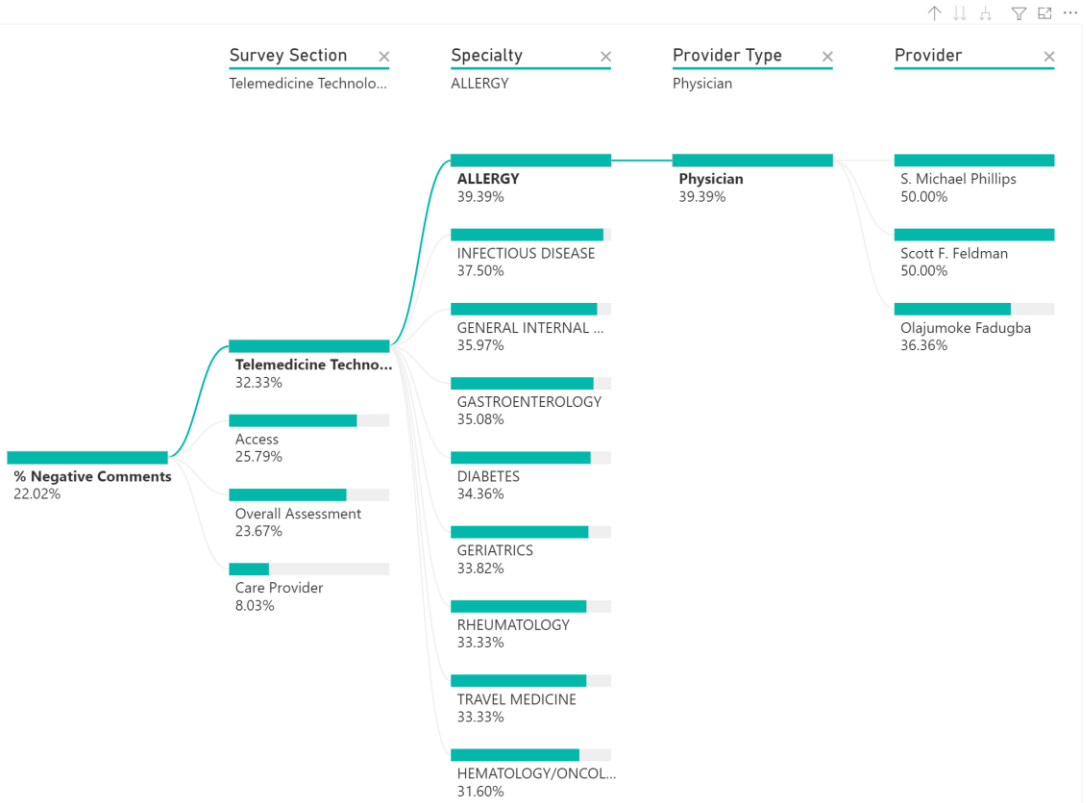
Press Ganey Sentiment Analysis: Current State

- DoM's current analysis product looks at Press Ganey comments using a word cloud format:
  - Provides a trending overview of the frequency of comments occur by affect (pos, neg, neut).



Press Ganey Sentiment Analysis: Current State (cont.)

- Root Cause Analysis Component:
  - Aims to answer the question: “Where are negative comments concentrated?”



IT MODE	Mixed	Negative	Neutral	Positive	Total
VIDEO	**811**	**2037**	**613**	**5821**	**9282**
Telemedicine Technology	209	679	224	928	**2040**
Overall Assessment	168	389	158	899	**1614**
Care Provider	141	205	91	2197	**2634**
Access	293	764	140	1797	**2994**
TELEPHONE	**232**	**710**	**511**	**1738**	**3191**
Telemedicine Technology	44	238	232	282	**796**
Overall Assessment	53	138	119	302	**612**
Care Provider	56	70	65	600	**791**
Access	79	264	95	554	**992**
Total	**1043**	**2747**	**1124**	**7559**	**12473**

Press Ganey Sentiment Analysis: Current State (cont.)

- Limitations of the current analysis products:
  - Overview
    - The high level summary doesn't immediately point to where or with whom the negative comments are being generated. It also does not indicate what the subject matter of the comments.
  - Root Cause
    - % of negative comments is based on frequency of occurrences
      - Providers that have 1 negative comment and no others will be ranked higher (100% neg) than a provider with multiple negative comments where that provider has 4 negative comments out of 10 comments (40%).

Natural Language Processing: VADER

Background

- Sentiment analysis is the process of determining affect behind a sentence or statement.
  - The majority of sentiment analysis approaches take one of two forms: **polarity-based**, where pieces of texts are classified as either positive or negative, or **valence-based,** where the intensity of the sentiment is taken into account.
  - Department of Medicine has historically only used a polarity based approach.

VADER (Valence Aware Dictionary and sEntiment Reasoner)

- Developed by C.J. Hutto and Eric Gilbert at Georgia Tech
- Is a lexicon and rule-based sentiment analysis tool that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion.
  - VADER sentimental analysis relies on a **sentiment lexicon**; a dictionary that maps lexical features to emotion intensities known as sentiment scores.

Word	Sentiment Score
tragedy	-3.4
rejoiced	2.0
insane	-1.7
disaster	-3.1
great	3.1

VADER Mechanics

- The toolkit breaks down full sentences and applies sentiment scores to each word in the sentence while accounting for contextual meaning.

- Lexical dictionary also accounts for:
  - Emoticons ☹
  - Acronyms (smh)
  - Emphasis from capitalizations and punctuation

```
print_sentiment_scores("The food is good.")
```

```
The food is good.----------------------- {'neg': 0.0, 'neu': 0.508, 'pos': 0.492, 'compound': 0.4404}
```

```
print_sentiment_scores("The food is GOOD!")
```

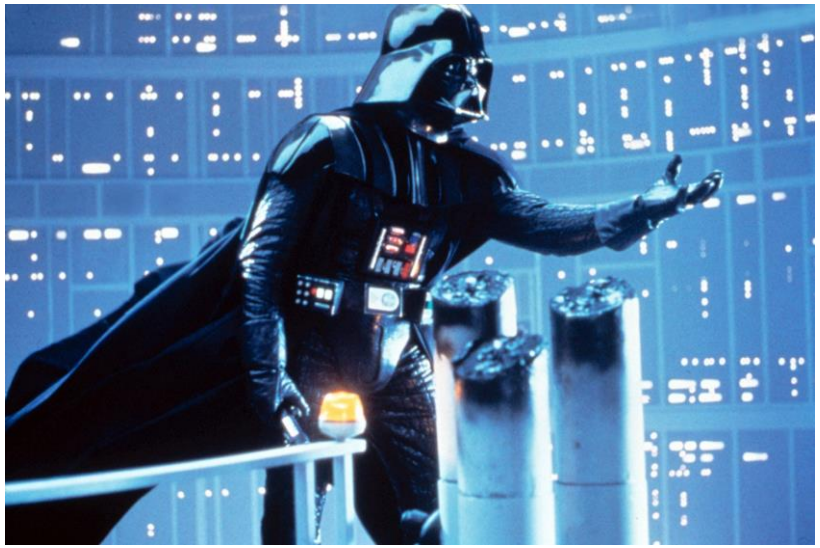
```
The food is GOOD!----------------------- {'neg': 0.0, 'neu': 0.433, 'pos': 0.567, 'compound': 0.6027}
```

VADER Mechanics Cont.

- Four sentiment metrics are produced:
  - Pos/Neu/Neg score, and a Compounded score (normalized between -1 and 1)

	Positive	Neutral	Negative	Compound
a couple of moments when audio was garbled, but video wasn't lost. did not impair conversation.	0.149	0.851	0	0.3491
useless for my condition. infuriating.	0	0.326	0.674	-0.7351
great doctor listens to you in very carefully great at Explaining diagnosisVery caring and easy to talk to	0.553	0.447	0	0.9441

- Allows analysts to numerically quantify the intensity of a Press Ganey comment, giving us insight into where patients feel the strongest about issues related to their appointment.



References

http://comp.social.gatech.edu/papers/icwsm14.vader.hutto.pdf

https://towardsdatascience.com/word-vectors-and-semantics-2863e7e55417

https://medium.com/analytics-vidhya/simplifying-social-media-sentiment-analysis-using-vader-in-python-f9e6ec6fc52f

https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664