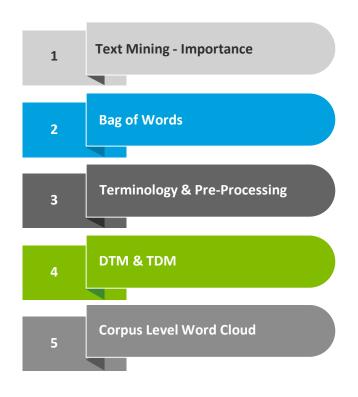
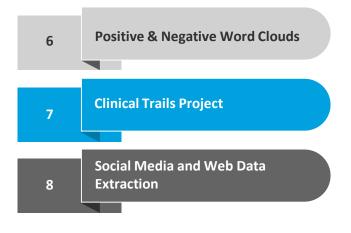


# Text Mining & Word Cloud

# Agenda







# Word Cloud Example

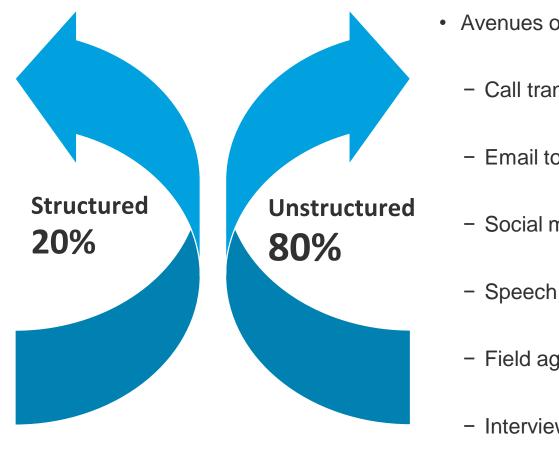






# Text Mining - Importance





- Avenues of textual unstructured data
  - Call transcripts



- Email to customer service



Social media outreach



- Speech transcripts



- Field agents, salespeople



Interviews & surveys



# Bag of Words





All the world's a stage, and all the men and women merely players:
They have their exits and their entrances;
And one man in his time plays many parts..."



World	Stage	Men	Women	Play	Exit	Entrance	time
1	1	2	1	2	1	1	1

# Terminology & Pre-Processing



• Each row is called as a 'Document' & even an empty row is considered as a document



• Collection of all these documents is called as 'Corpus' / 'Corpora'







- Quirks of languages
  - Terms with typos (e.g., 'musc')
  - Terms in lowercase, proper case & uppercase (e.g., usb, Usb, USB)
  - Punctuations & special symbols ('%', '!', '&', etc.)
  - Filler words, connectors, pronouns ('all', 'for', 'of', 'my', 'to', etc.)
- Stemming process of considering only stem words (e.g., jumping, jumped; stem-word here is 'jump')

### Let me show you "Amazon customer reviews"

### DTM & TDM







• Let us understand 100 document corpus of Xbox

**DTM** weighing

**TF - Regular term counts** 

**TFIDF - Discounts the TF by document frequency** 

	DTM with TF weighing							
Documents	Terms ->							
Documents	Keyboard	Usb	Free	Hard				
1	4	1	0	2				
2	3	1	1	0				
•••				•••				
50	2	3	0	1				
				•••				
70	1	2	0	3				
		•••	•••	•••				
100	0	0	1	0				
Terms Sum	200	50	100	80				
Doc. Freq (DF)	2	0.5	1	0.75				

# Corpus Level Word Cloud





### **Positive Word Cloud**



```
affordable
                                recommended
prefer
                 available fantastic
                                              greatest
 gladly
                                        delighted smile
                                         blissfully ready
clean
friendly
   cheaper
   improvement Worked fans
       complement compatible delightful enjoyed
honor
```

## **Negative Word Cloud**











www.360digitmg.com



info@360digitmg.com