SLIDE-1

Clinical Trials - Text Analytics

SLIDE-2

Before we begin…..

The text from clinical trials has been analyzed using various business analytics techniques such as Word Cloud using Term Frequency (TF) & TF Inverse Document Frequency (TFIDF) including unigram, bigram & trigram, Sentiment analysis, Semantic networks, Cluster Dendrogram, Unsupervised learning (K-Means clustering), Association rules etc

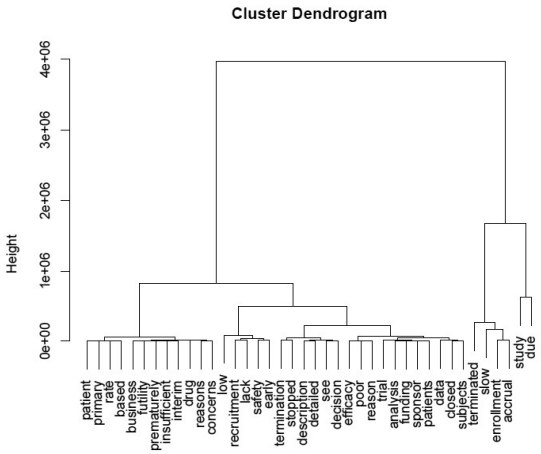
SLIDE-3

- Word Cloud – Unigram, Cluster Dendrogram
- Semantic network, Bigram
- K – Means clustering – Scree-plot
- K – Means clustering – Cluster 1
- Association Rules – Report yet to be prepared, work completed
- Prediction model – Not yet started working
- Numerical analysis from detailed description – Not yet started working

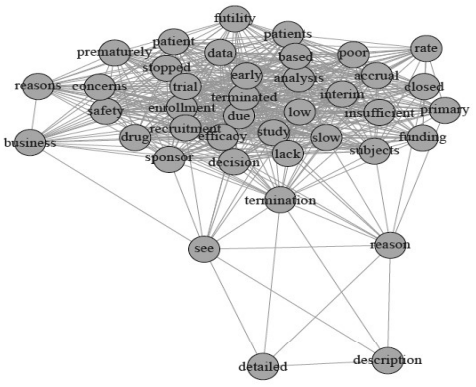
SLIDE-4

- Key words standing out of the rest are Accrual, Enrollment, Slow, Safety, Efficacy, Sponsor, Lack, Low etc.
- These words should be seen in the context to gain business value
- When we see this word cloud in conjunction with dendrogram, we notice that slow accrual, slow enrollment, poor efficacy, sponsor

funding seem to be the broad themes for termination of clinical trials



SLIDE-5

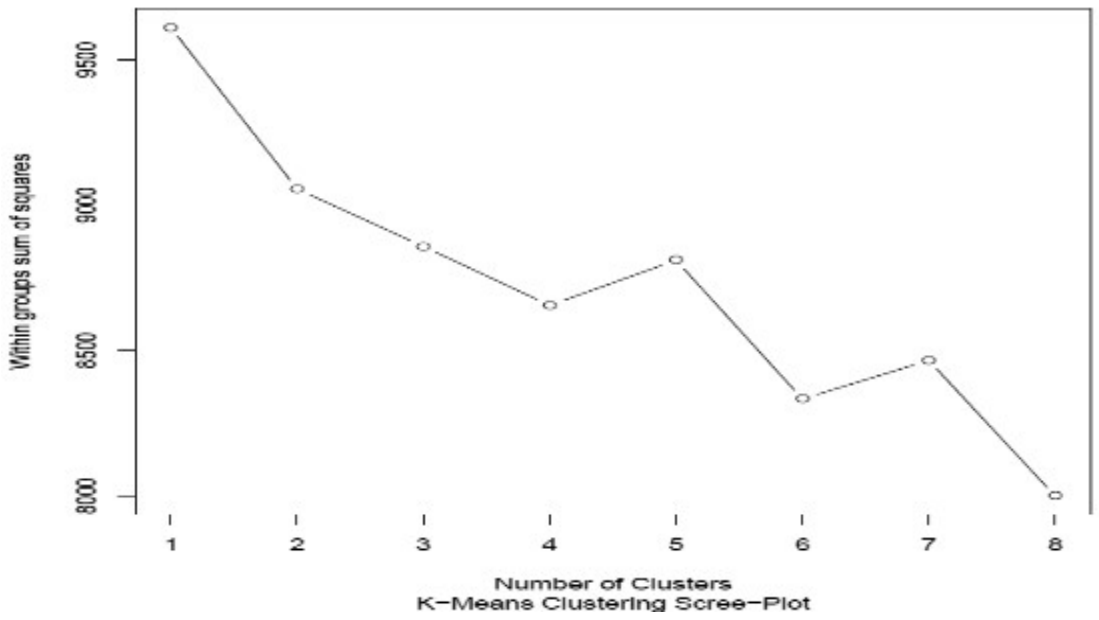


- Semantic network shows that the relationship between the words & the key themes mentioned in previous slide are becoming relevant
- One key thing is safety concerns. At the first sight it sounds as if safety concerns were reason for termination, but when we see it

in context, more termination reasons say that there are “No Safety Concerns”

- Bigram is used to see 2 words to extract business value & the key themes mentioned earlier are more evident here

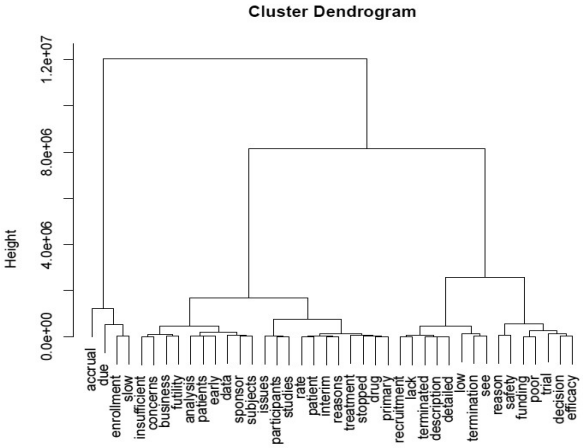
SLIDE-6



- Scree-plot or elbow plot shows that there is a clear bend at 2 clusters, hence we are considering that there are 2 clusters (categories) that the data can be segregated into

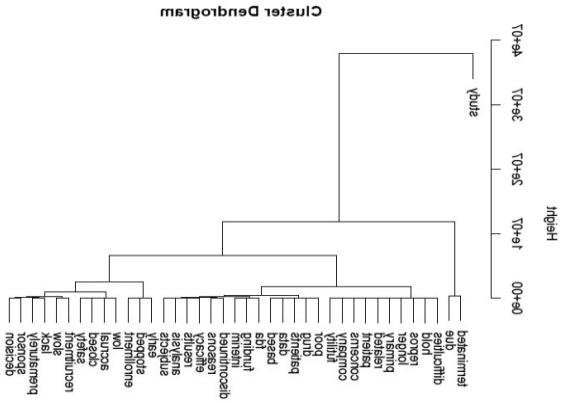
Note: Analysis is done considering slight bend at 2nd cluster and considering steep bend at 4th cluster, however, it did not provide any meaningful insights

SLIDE-7



- Word cloud is clearly highlighting that this cluster is speaking majorly about Accrual:
  Term referring to the number of patients in a study or clinical trial
- Even the dendrogram clearly shows Accrual, Enrollment, Slow as a major cluster

SLIDE-8



- Few key highlights from this word cloud are early & premature termination
- Dendrogram mentions majority of things related to premature closure of clinical trials