



**THE
POWER
TO KNOW.®**

Turbocharge Your Analytics with Text Mining

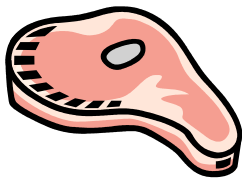
SAS Melbourne User Group
December 10, 2008

Challenge of Natural Language Processing

- English Sentence

“Flesh is weak but the spirit is strong”

- Russian Translation



“Meat is rotten but the alcohol is great!”

Agenda

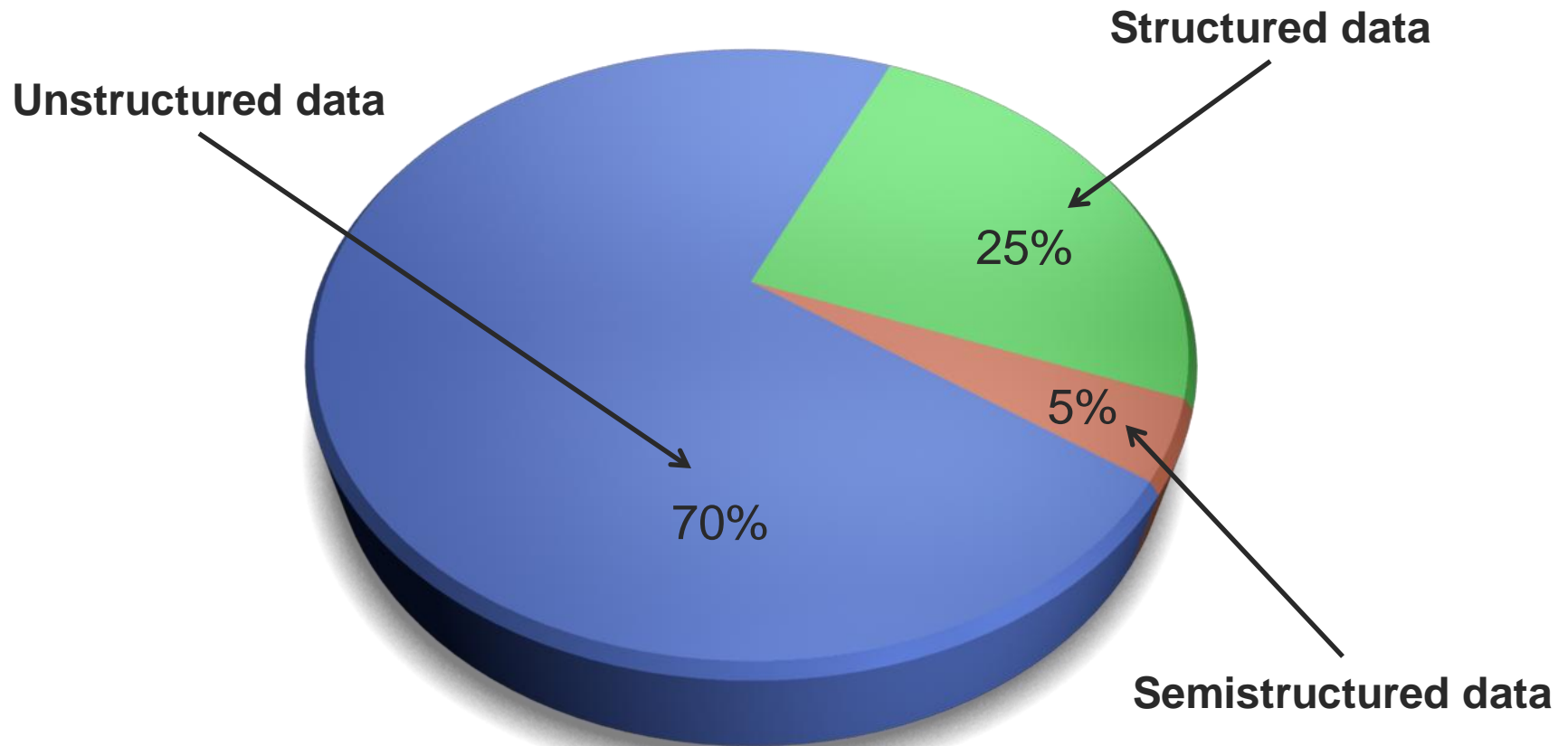
- Enterprise Information in unstructured form
- Business Need for Text Analytics
- SAS Text Miner
- Demo
- Q&A

Enterprise Information

- Unstructured data comes in many forms
- Doubling at an unprecedented rate
- Left unmanaged will decay
- Managed and invested in silo's

Unstructured and Semi-structured Data

The Dark Matter for IT



Agenda

- Enterprise Information in unstructured form
- Business Need for Text Analytics
- SAS Text Miner
- Demo
- Q&A

Business Need for Text Analytics

Government

- Early warning of terrorist behavior
- Crime ring and criminal tracking
- Customs and border control
- Healthcare fraud
- Anti money laundering
- Tax evasion/fraud
- Social Network Analysis

Business Need for Text Analytics

Industry

- Finance, Insurance, Life Sciences, Retail, Telco
 - Identify most profitable customers
 - Improve brand imaging
 - Prevent churn
 - Identify adverse drug reactions
 - Help boost cross-sell/up-sell opportunities
 - Early warning for fraudulent claims

Text Analytics Techniques for Unstructured Data

OBS #	Symptoms Text
5	Cellulitus at administration site.
7	Demyelinating disease; dizziness, blurred vision; difficulty hearing and walking.
11	Autistic mannerisms, system "shutdown". Blank stares, catatonic state.
12	Loss of speech and coordination.
27	Patient experienced a seizure
73	Rt arm injured; movement impaired.
86	The pt developed erythema multiforme within 48 hours, post vax.
88	11 by 16 indurated erythematous warm area right upper thigh on 12/15/2001.
92	Fever up to 101F, irritability and erythema at the site.

Agenda

- Enterprise Information in unstructured form
- Business Need for Text Analytics
- SAS Text Miner
- Demo
- Q&A

Proactive Decision Making with SAS Text Miner

Features

- Text parsing
- Part-of-speech tagging
- Stemming
- Phrase Recognition
- Customize start/stop lists, synonyms, abbreviations etc.
- Advanced Visualizations

Proactive Decision Making with SAS Text Miner

Benefits

- Full integration with SAS Enterprise Miner's data analytics
- Easily combine structured and unstructured data
- Broad set of tools to support the complete text mining process
- An open, extensible design for ultimate flexibility
- Build more and better models faster!

Agenda

- Enterprise Information in unstructured form
- Business Need for Text Analytics
- SAS Text Miner
- Demo
- Q&A

Scenario– Vaccine Adverse Reactions

- Department of Health and Human Services
- Vaccine adverse reaction as symptoms text
- Providers required to report contraindications and reactions
- Text contains medical terms, abbreviations, misspellings

- “The pt developed itchiness, rubor and an erythematous lesion at the site of the injection. Treated with 250mg Solu-Medrol injection (IM)”

Text Mining Clusters

Text Miner - Interactive

File Edit Tools View Window

Clusters

#	Descriptive Terms	Freq	Percentage	RMS Std.
1	+ month, + old, + hepatitis, medical, + report, + year, male, + female, + experience, + time, + dose, information	1492	0.141609719...	0.1544128...
2	+ site, + virus, + experience, + year, + health, + treatment, concomitant, + month, + injection site, + report, + old, + injection	1186	0.112566438...	0.1231781...
3	+ last, + episode, + hour, + week, + minute, + headache, + feel, + fever, + cry, + temperature, + vomit, + symptom	976	0.092634776...	0.1225332...
4	+ leave, + pain, + leave arm, + deltoid, + arm, + right, + week, + feel, + female, + year, + shoot, + dose	1203	0.114179954...	0.1323010...

Agenda

- Enterprise Information in unstructured form
- Business Need for Text Analytics
- SAS Text Miner
- Demo
- Q&A



**THE
POWER
TO KNOW®**