Textual Analytics for Accounting and Auditing
Thanks to

- Ingrid Fisher (SUNY – Albany)
  - Research interests - textual analysis in accounting, design science in accounting and financial accounting standards/documents
- Jordan Seebach (Grant Thornton)
  - Audit Manager
  - 6 years with Grant Thornton
  - 18 month rotation in the Audit Methodology and Standards Group
  - Data Analytics Champion
- Lorraine Lee (UNC- Wilmington)
Lab access

• Account name: ciia2018
• Password: Accounting
• Files needed can be found at:
  • C:\Users\Public\Documents
Outline

- Motivation
- Textual Analysis Definition
- How Textual Analysis is Used in Accounting
- Textual Analysis Key Terms
- Textual Analysis Methods
- Practice
- Summary
Motivation

- Accountants prepare complex accounting footnotes
- Managers prepare MDA
- Do investors / regulators really want to read through all these complex footnotes and MDA to answer specific questions such as
  - What are the firm’s new products?
  - What are the details of its lease obligations?
  - What are the details of the firm’s contingent liabilities?
Motivation

- Large corporations enter into multiple lease contracts
- New revenue recognition standard requires regular review of customer contracts
- Can accountants / auditors manually review these lease contracts and customer contracts to ensure proper lease accounting and revenue recognition?
Motivation

• Auditors need to sort through millions of client journal entries
• Each journal entry includes account information and entry description
• Can auditors identify transactions to investigate further based upon a review of journal entry descriptions?
Motivation

- PCAOB inspects all Big 4 and Second Tier audit firms annually and makes publicly available its inspection reports.
- Inspection reports describe specific problems auditors missed for selected clients (issuers per PCAOB).
- General public / investors / board of directors want to know:
  - What problems did the auditors miss?
  - Did these problems result in material misstatements to issuers?
  - Have these problems been resolved or are auditors continuing to incur the same problems?
Textual Analysis Definition

- A systematic analysis of the content rather than the structure of a communication, such as a written work, speech, or film, including the study of thematic and symbolic elements to determine the objective or meaning of the communication. (thefreedictionary.com)
- Synonym – context analysis, text mining, data mining
- Based on linguistics theory
Why Textual Analysis Now?

- Exponential increase in computing power over past two decades
- Increased focus on textual methods driven by requirements of internet search engines
- Technique has permeated most disciplines in one way or another
- In accounting and finance, online availability of news articles, earnings conference calls, Securities and Exchange Commission (SEC) filings, and text from social media provide ample fodder for applying textual analysis technology
Lanza Approach to Letter Analytics

Identifies word deviations swiftly by relating letter frequency patterns to benchmarks of the English language and prior period letter occurrences

- Focuses on first letter, last letter, first two letters, last two letters
- Benchmarked against prior periods
- Benchmarking against peer group or industry is not effective
- Analyzing actual words as well as meaning,
Lanza Approach to Letter Analytics

- Used in the risk assessment and planning phases of the audit
- Types of data analyzed
  - General ledger data through journal entry descriptions
  - Public filings
  - Earnings calls transcripts
  - News articles
- MD&A of 10-K and 10-Q has most meaningful data
Lanza Approach to Letter Analytics

- What do the analytics identify?
  - Transactions that are unique to the period
  - Change in words used in journal entry descriptions
  - Change in tone of wording in public filings
  - Tendencies in management statements

- Business use
  - Process and transaction flow insight
  - Journal entry trends
  - Profile employees for corruption or collusion
  - Pinpoint computer application issues or concerns
General Ledger Fingerprint
Examples of Fraud Cases

- WorldCom
  - Capitalizing interconnection expenses with other telecom companies
  - Inflating revenues with corporate unallocated revenue accounts
- HealthSouth
  - Over 100,000 entries each month to capitalize amounts under $5,000
  - Capitalized journal entry description of all fraudulent entries
Contract Analysis

- Ability to analyze all contracts for a given company
- Identify differences between existing contracts and standard contracts
- Calculates percentage of consistency with standard
- Helps teams identify:
  - Unique items that have accounting implications
  - Embedded leases
  - Embedded derivatives
Expectations of accountants

Characteristics that support the analytical approach

• Focused on process improvement/challenging the norm
• Drive efficiencies in processes
• Understand how to analyze data
  • "Cleaning up" data
  • Normalizing data
What textual analysis related skills do accountants need?

- Understanding of basic textual analysis process and vocabulary
- When to use textual analysis vs another technique
- Advantages of using textual analysis
- What questions to ask when evaluating textual analysis results?
Textual Analysis Key Terms

- Word count
- Word cloud
- Word tree
- Word search
- Fog index / readability
- Tone / sentiment analysis
Textual Analysis Key Terms – Word Count

- **Count** number of words in document
- Can also count number of pages, paragraphs, and lines in your document
- Can also display number of characters, either including or excluding spaces
Textual Analysis Key Terms – Word Cloud

- graphical representations of word frequency that give greater prominence to words that appear more frequently in a source text
- larger the word in visual the more common the word was in the document(s)
- type of visualization to assist evaluators by identifying words that frequently appear in set of interviews, documents, or other text.
- can also be used to communicate most salient points or themes in reporting stage
Textual Analysis Key Terms – Word Cloud

Motivation  Definition  Uses  Key Terms  Methods  Practice
Textual Analysis Key Terms – Word Tree

- Show pre-selected word(s) and how it is connected to other words in text-based data through visual branching structure.
- Unlike word clouds, word trees visually display connection of words in dataset, providing some context to their use.
Textual Analysis Key Terms – Word Tree

Firm failed to perform sufficient financial statements or reporting on a level of reliance and placed excessive auditing revenue placed on the effectiveness of ICFR. review a service auditor’s administrator’s / or on statements and on the effectiveness of internal control.

10. Issue) In this and substantive testing with respect below. • The Firm excluded the in its approach to testing or otherwise test the relevant ("ICFR"). In addition, audit in Specifically The over financial reporting ("ICFR"). Specifically, The The Firm that was excessive specifically, the due to With respect to the opinion

Iowa State University
Textual Analysis Key Terms – Word Search

- Most frequent phrases and frequencies of words
- Many support non-English language texts
- Can be used to analysis content
- Can provide lexical density – i.e. number of **lexical** words (or content words) divided by the total number of words
- **Lexical** words give text its meaning and provide information regarding what text is about.
- More precisely, **lexical** words are simply nouns, adjectives, verbs, and adverbs
Textual Analysis Key Terms – 
Fog Index / Readability

- Tests are designed to indicate how difficult a passage in English is to understand
- Also labeled Gunning-Fog index
- Linear combination of average sentence length and proportion of complex words (words with more than two syllables)
Textual Analysis Key Terms – Fog Index / Readability

• Two tests
  • Flesch Reading Ease
  • Flesch–Kincaid Grade Level

• Tests have same core measures (word length and sentence length)

• Tests use different weighting factors

• Results of two tests correlate approximately inversely
  • Text with comparatively high score on Reading Ease test should have lower score on Grade-Level test
Textual Analysis Key Terms – Fog Index / Readability – The Financial Statement Challenge

- Former SEC chair Christopher Cox suggested that Fog Index can gauge compliance with the SEC’s plain English initiative (1998)
- Research shows Fog index is not a good measure of financial statement readability (Loughran and McDonald 2014)
- Why not?
  - Business text has an extremely high % of complex words that are generally understood by investors and analysts
  - File size of 10K can proxy for document readability
Your Turn - Textual Analysis Key Terms – Fog Index – The Financial Statement Challenge

• What complex words do you think might appear in 10K filings?
Textual Analysis Key Terms – Fog Index / Readability for Financial Statements

- Loughran and McDonald (2014, 2017) developed a readability scale for financial statement analysis
- Key words include:
Textual Analysis Key Terms - Tone

- Positive
- Negative
- Ambiguous
  - use of uncertain (e.g., approximate, contingency, uncertain, and indefinite) and weak modal words (e.g., might, possible, approximate, and contingent)
- Challenging – compare use of ‘call’ in:
  - ‘Firm A grant call options to managers’
  - ‘Firm A call back inferior products’
Three Steps to Textual Analysis

- Harvest text
- Clean and parse text
- Analyze text
Harvest Data

• Collect *data* from forum or web, such as Yahoo finance forum and twitter
• Alternatively, collect *data* from financial database, such as Thomson Reuter News Database, newspaper databases
• File format varies
  - Txt
  - Xml
  - Pdf
Sources of Unstructured Data to Examine

• Mandatory filings and disclosures (e.g., 10-Ks, 10-Qs, 8-Ks annual reports, IPO prospectuses, RNS, etc.)
• Earning announcements and other press releases
• Conference calls (management presentation and Q&A sections) and investor road show presentations
• Financial media articles (e.g. WSJ, DJNS, FT, newswire service, etc.)
Sources of Unstructured Data to Examine

• Analyst reports and research notes
• Regulatory announcements (e.g., SEC litigation releases)
• Macro and sentiment news (e.g., Federal Open Market Committee minutes)
• Internet message boards
• Social networks (e.g. Seeking Alpha [http://seekingalpha.com](http://seekingalpha.com))
Your Turn - Sources of Unstructured Data to Examine

• What sources of unstructured data could you examine to address relevant business problems?
Clean and Parse Data

- Using unstructured data
- Remove taggers and stop words thus putting plain text into a word vector

- **Your Turn** – what words should be removed??
Analyze Data

- Many different techniques available to use
- Some require both manual and computerized interventions
Textual Analysis Methods

- **Machine Learning**
  - Text analysis often relies on machine learning, a branch of computer science that trains computers to recognize patterns.
  - There are two kinds of machine learning used in text analysis:
    - supervised learning, where a human helps to train pattern-detecting model – Naïve Bayes Classification
    - unsupervised learning, where computer finds patterns in text with little human intervention - Natural Language Processing and Topic Modeling
Textual Analysis Methods

- Natural Language Processing
  
  Natural language processing, kind of machine learning, is attempt to use computational methods to extract meaning from free text. Among other things, natural language processing algorithms can derive:
  
  - names of people and places
  - dates
  - sentiment
  - parts of speech
Textual Analysis Methods

- **Topic Modeling**
  - Topic modeling, a form of machine learning, is a way of identifying patterns and themes in a body of text
  - Topic modeling is done by statistical algorithms, such as Latent Dirichlet Allocation, which groups words into "topics" based on which words frequently co-occur in a text
Textual Analysis Methods

- **Network Analysis**
  - Network analysis is a method for finding connections between nodes representing people, concepts, sources, and more.
  - These networks are usually visualized into graphs that show the interconnectedness of the nodes.
Textual Analysis Methods – Network Analysis
Textual Analysis Methods

• Citation Analysis
  • Like network analysis, this research method can be used to discover connections and relationships between various citations of documents and then visualized
Example – Used textual analysis to identify information technology control deficiencies in PCAOB inspection reports

- Evaluated 48 PCAOB inspection reports from inspection years 2010 to 2015
- Included all Big 4 and Second Tier auditing firms (Grant Thorton, RSM, BDO, Crowe Howarth)
- Looked for deficiencies related to control deficiencies
- Classified as entity-level vs application-level control deficiencies
- Found more application-level control deficiencies
- Discovered approximately same number of deficiencies in inspection year 2015 as in inspection year 2010
Example: Does gender diversity in the audit committee influence key audit matters’ readability in the audit report? UK evidence.

- Forthcoming in *Corporate Social Responsibility and Environmental Management* by Dr. Patrick Velte
- Looks at relationship between percentage of women on audit committees in UK firms and auditors’ disclosure of key audit matters (KAM) in 2014 and 2015
- Find that UK companies with higher percentage of women on audit committees are more likely to have higher readability of KAM disclosures as measured by Flesch reading ease index
- Results hold for Fog readability index and Blau index
Use in Accounting – Internal Decision Making

- New revenue recognition standards
- New leasing standard
- Social media sentiment
Use in Accounting – External Decision Making

- Review corporate annual reports for investment decision making
- Review additional information corporations provide for decision making
- Social media sentiment
Questions for Textual Analysis

• Can we tease out sentiment from mandated company disclosures and contextualize quantitative data in ways that might predict future valuation components?

• Can we computationally read news articles and trade before humans can read and assimilate the information?

• If Twitter’s tweets provide the pulse of information, can we monitor these messages in real time to gain an informational edge?

• Do textual artifacts provide an additional attribute that predicts bankruptcies?
Questions for Textual Analysis

• Are there subtle cues in managements’ earnings conference calls that computers can discern better than analysts?
• More broadly, can we examine textual artifacts to measure the quantity and quality of information in a collection of text, including both intended message and, importantly, any unintended revelations?
Your Turn – What Questions Could Textual Analysis Help With?

- How do you think you might be able to use textual analysis in your job (or personal life)?
Challenges learning textual analysis skills

• What task to use to demonstrate textual analysis software?
• Where can I find the data needed?
• Cost of textual analysis software
The activity - DETAILS

- Discuss
  - how textual analytics in accounting and business (in general) has grown in popularity
  - how textual analysis is an important tool for mining unstructured data
- Expose participants to textual analysis
  - Use publicly available data
  - Use free textual analysis software (i.e. Rapid Miner)
Open RapidMiner and Acquire Appropriate Extensions

Choose a template to start from:

- **Blank**
  - Start with a blank process.

- **Churn Modeling**
  - Predict which of your customers will churn and why with a decision tree.

- **Direct Marketing**
  - Predict response to campaigns and increase the conversion rate of your campaign.

- **Credit Risk Modeling**
  - Model credit default risk by training an optimized Support Vector Machine (SVM) model.

- **Market Basket Analysis**
  - Find products frequently purchased together and turn them into rules for recommendations.

- **Predictive Maintenance**
  - Model equipment failures to schedule maintenance preemptively.

- **Price Risk Clustering**
  - Cluster price developments using X-Means to unveil price-risk relationships.

- **Lift Chart**
  - Create a lift chart to visualize the improvement that a model provides compared to guessing.
Toggle to Start tab and choose blank
Home Page of RapidMiner

Your process looks empty. Add some data first. Drag data or operators here.

Leverage the Wisdom of Crowds to get operator recommendations based on your process design!

Activate Wisdom of Crowds

Motivation  Definition  Uses  Key Terms  Methods  Practice
Add Appropriate Extensions to Analyze Text, download “text processor” extension by selecting extensions along top border.
Under marketplace, select text processing

Need to also add AYLIEN Text Analysis
Tokenize shows list of words in 10-Q
## Initial Word List

<table>
<thead>
<tr>
<th>Word</th>
<th>Attrib...</th>
<th>Tot...</th>
<th>Docum...</th>
<th>Wells F...</th>
<th>JPMorg...</th>
<th>Bank o...</th>
</tr>
</thead>
<tbody>
<tr>
<td>the</td>
<td>the</td>
<td>9801</td>
<td>3</td>
<td>3161</td>
<td>3639</td>
<td>3001</td>
</tr>
<tr>
<td>and</td>
<td>and</td>
<td>9362</td>
<td>3</td>
<td>3121</td>
<td>2938</td>
<td>3303</td>
</tr>
<tr>
<td>of</td>
<td>of</td>
<td>7728</td>
<td>3</td>
<td>2752</td>
<td>2662</td>
<td>2314</td>
</tr>
<tr>
<td>in</td>
<td>in</td>
<td>4833</td>
<td>3</td>
<td>1800</td>
<td>1619</td>
<td>1414</td>
</tr>
<tr>
<td>to</td>
<td>to</td>
<td>4647</td>
<td>3</td>
<td>1649</td>
<td>1432</td>
<td>1566</td>
</tr>
<tr>
<td>for</td>
<td>for</td>
<td>2929</td>
<td>3</td>
<td>1063</td>
<td>957</td>
<td>909</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>2773</td>
<td>6</td>
<td>870</td>
<td>1177</td>
<td>726</td>
</tr>
<tr>
<td>loans</td>
<td>loans</td>
<td>2208</td>
<td>3</td>
<td>680</td>
<td>752</td>
<td>776</td>
</tr>
<tr>
<td>on</td>
<td>on</td>
<td>2096</td>
<td>3</td>
<td>652</td>
<td>756</td>
<td>688</td>
</tr>
<tr>
<td>Septem...</td>
<td>Septem...</td>
<td>2031</td>
<td>3</td>
<td>450</td>
<td>656</td>
<td>925</td>
</tr>
<tr>
<td>Total</td>
<td>Total</td>
<td>1562</td>
<td>3</td>
<td>586</td>
<td>527</td>
<td>449</td>
</tr>
<tr>
<td>The</td>
<td>The</td>
<td>1537</td>
<td>3</td>
<td>424</td>
<td>669</td>
<td>444</td>
</tr>
<tr>
<td>or</td>
<td>or</td>
<td>1522</td>
<td>3</td>
<td>570</td>
<td>510</td>
<td>442</td>
</tr>
<tr>
<td>as</td>
<td>as</td>
<td>1508</td>
<td>3</td>
<td>504</td>
<td>525</td>
<td>479</td>
</tr>
<tr>
<td>securities</td>
<td>securities</td>
<td>1495</td>
<td>3</td>
<td>667</td>
<td>405</td>
<td>423</td>
</tr>
<tr>
<td>by</td>
<td>by</td>
<td>1440</td>
<td>3</td>
<td>478</td>
<td>566</td>
<td>396</td>
</tr>
<tr>
<td>credit</td>
<td>credit</td>
<td>1424</td>
<td>3</td>
<td>389</td>
<td>544</td>
<td>491</td>
</tr>
<tr>
<td>million</td>
<td>million</td>
<td>1332</td>
<td>3</td>
<td>387</td>
<td>240</td>
<td>705</td>
</tr>
</tbody>
</table>
Filter Tokens (by length) and Filter Stopwords (English)

<table>
<thead>
<tr>
<th>Word</th>
<th>Attributed</th>
<th>Total</th>
<th>Document</th>
<th>Wells F.</th>
<th>JPMorgan</th>
<th>Bank of</th>
</tr>
</thead>
<tbody>
<tr>
<td>loans</td>
<td>loans</td>
<td>2604</td>
<td>3</td>
<td>819</td>
<td>865</td>
<td>920</td>
</tr>
<tr>
<td>septem...</td>
<td>septem...</td>
<td>2031</td>
<td>3</td>
<td>450</td>
<td>656</td>
<td>925</td>
</tr>
<tr>
<td>total</td>
<td>total</td>
<td>1896</td>
<td>3</td>
<td>727</td>
<td>604</td>
<td>565</td>
</tr>
<tr>
<td>credit</td>
<td>credit</td>
<td>1884</td>
<td>3</td>
<td>545</td>
<td>668</td>
<td>671</td>
</tr>
<tr>
<td>securities</td>
<td>securities</td>
<td>1814</td>
<td>3</td>
<td>808</td>
<td>526</td>
<td>480</td>
</tr>
<tr>
<td>net</td>
<td>net</td>
<td>1586</td>
<td>3</td>
<td>525</td>
<td>525</td>
<td>536</td>
</tr>
<tr>
<td>assets</td>
<td>assets</td>
<td>1405</td>
<td>3</td>
<td>566</td>
<td>399</td>
<td>440</td>
</tr>
<tr>
<td>value</td>
<td>value</td>
<td>1382</td>
<td>3</td>
<td>485</td>
<td>425</td>
<td>472</td>
</tr>
<tr>
<td>million</td>
<td>million</td>
<td>1334</td>
<td>3</td>
<td>387</td>
<td>240</td>
<td>707</td>
</tr>
<tr>
<td>billion</td>
<td>billion</td>
<td>1308</td>
<td>3</td>
<td>439</td>
<td>294</td>
<td>575</td>
</tr>
<tr>
<td>income</td>
<td>income</td>
<td>1244</td>
<td>3</td>
<td>423</td>
<td>418</td>
<td>403</td>
</tr>
<tr>
<td>interest</td>
<td>interest</td>
<td>1121</td>
<td>3</td>
<td>376</td>
<td>354</td>
<td>391</td>
</tr>
<tr>
<td>risk</td>
<td>risk</td>
<td>1097</td>
<td>3</td>
<td>421</td>
<td>363</td>
<td>313</td>
</tr>
<tr>
<td>months</td>
<td>months</td>
<td>1083</td>
<td>3</td>
<td>255</td>
<td>361</td>
<td>467</td>
</tr>
<tr>
<td>fair</td>
<td>fair</td>
<td>1032</td>
<td>3</td>
<td>354</td>
<td>344</td>
<td>334</td>
</tr>
<tr>
<td>losses</td>
<td>losses</td>
<td>1028</td>
<td>3</td>
<td>336</td>
<td>378</td>
<td>314</td>
</tr>
<tr>
<td>mortgage</td>
<td>mortgage</td>
<td>1014</td>
<td>3</td>
<td>506</td>
<td>246</td>
<td>262</td>
</tr>
<tr>
<td>loan</td>
<td>loan</td>
<td>1005</td>
<td>3</td>
<td>306</td>
<td>344</td>
<td>355</td>
</tr>
</tbody>
</table>
## Passive Words

<table>
<thead>
<tr>
<th>Word</th>
<th>Attribute</th>
<th>Total O..</th>
<th>Docum...</th>
<th>Well...</th>
<th>JPMorg...</th>
<th>Bank o...</th>
</tr>
</thead>
<tbody>
<tr>
<td>related</td>
<td>related</td>
<td>777</td>
<td>3</td>
<td>216</td>
<td>354</td>
<td>207</td>
</tr>
<tr>
<td>based</td>
<td>based</td>
<td>457</td>
<td>3</td>
<td>172</td>
<td>180</td>
<td>105</td>
</tr>
<tr>
<td>ended</td>
<td>ended</td>
<td>984</td>
<td>3</td>
<td>172</td>
<td>364</td>
<td>448</td>
</tr>
<tr>
<td>held</td>
<td>held</td>
<td>439</td>
<td>3</td>
<td>158</td>
<td>148</td>
<td>133</td>
</tr>
<tr>
<td>backed</td>
<td>backed</td>
<td>266</td>
<td>3</td>
<td>152</td>
<td>64</td>
<td>50</td>
</tr>
<tr>
<td>included</td>
<td>included</td>
<td>395</td>
<td>3</td>
<td>143</td>
<td>122</td>
<td>130</td>
</tr>
<tr>
<td>compared</td>
<td>compared</td>
<td>301</td>
<td>3</td>
<td>132</td>
<td>44</td>
<td>125</td>
</tr>
<tr>
<td>preferred</td>
<td>preferred</td>
<td>187</td>
<td>3</td>
<td>119</td>
<td>24</td>
<td>44</td>
</tr>
<tr>
<td>continued</td>
<td>continued</td>
<td>134</td>
<td>3</td>
<td>99</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>consolidated</td>
<td>consolidated</td>
<td>471</td>
<td>3</td>
<td>94</td>
<td>181</td>
<td>196</td>
</tr>
<tr>
<td>guarant...</td>
<td>guarant...</td>
<td>115</td>
<td>3</td>
<td>86</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>recorded</td>
<td>recorded</td>
<td>245</td>
<td>3</td>
<td>82</td>
<td>81</td>
<td>82</td>
</tr>
<tr>
<td>expected</td>
<td>expected</td>
<td>163</td>
<td>3</td>
<td>79</td>
<td>46</td>
<td>38</td>
</tr>
<tr>
<td>recogni...</td>
<td>recogni...</td>
<td>176</td>
<td>3</td>
<td>78</td>
<td>47</td>
<td>51</td>
</tr>
<tr>
<td>collater...</td>
<td>collater...</td>
<td>102</td>
<td>3</td>
<td>70</td>
<td>14</td>
<td>18</td>
</tr>
<tr>
<td>insured</td>
<td>insured</td>
<td>148</td>
<td>3</td>
<td>68</td>
<td>27</td>
<td>53</td>
</tr>
<tr>
<td>sold</td>
<td>sold</td>
<td>238</td>
<td>3</td>
<td>67</td>
<td>94</td>
<td>77</td>
</tr>
<tr>
<td>required</td>
<td>required</td>
<td>162</td>
<td>3</td>
<td>62</td>
<td>46</td>
<td>54</td>
</tr>
</tbody>
</table>
Readability

- Loughran and McDonald (2014, 2017) show that FOG index does not work well for financial statement disclosures
## Readability

<table>
<thead>
<tr>
<th>Word</th>
<th>Attribute</th>
<th>Total Occurrences</th>
<th>Document Frequency</th>
<th>Well Frequency</th>
<th>JP Morgan Frequency</th>
<th>Bank of America Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>risk</td>
<td>risk</td>
<td>1097</td>
<td>3</td>
<td>421</td>
<td>363</td>
<td>313</td>
</tr>
<tr>
<td>exposure</td>
<td>exposure</td>
<td>291</td>
<td>3</td>
<td>82</td>
<td>84</td>
<td>125</td>
</tr>
<tr>
<td>approx...</td>
<td>approx...</td>
<td>119</td>
<td>3</td>
<td>40</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>assump...</td>
<td>assump...</td>
<td>94</td>
<td>3</td>
<td>40</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>risks</td>
<td>risks</td>
<td>92</td>
<td>3</td>
<td>38</td>
<td>31</td>
<td>23</td>
</tr>
<tr>
<td>variable</td>
<td>variable</td>
<td>69</td>
<td>3</td>
<td>31</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>believe</td>
<td>believe</td>
<td>41</td>
<td>3</td>
<td>25</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>intangible</td>
<td>intangible</td>
<td>53</td>
<td>3</td>
<td>24</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>assump...</td>
<td>assump...</td>
<td>27</td>
<td>3</td>
<td>20</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>exposur...</td>
<td>exposur...</td>
<td>97</td>
<td>3</td>
<td>20</td>
<td>33</td>
<td>44</td>
</tr>
<tr>
<td>difference</td>
<td>difference</td>
<td>38</td>
<td>3</td>
<td>19</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>probable</td>
<td>probable</td>
<td>38</td>
<td>3</td>
<td>18</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>unobser...</td>
<td>unobser...</td>
<td>54</td>
<td>3</td>
<td>18</td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>contingent</td>
<td>contingent</td>
<td>28</td>
<td>3</td>
<td>15</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>intangib...</td>
<td>intangib...</td>
<td>32</td>
<td>3</td>
<td>15</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>differ</td>
<td>differ</td>
<td>33</td>
<td>3</td>
<td>14</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>indepen...</td>
<td>indepen...</td>
<td>21</td>
<td>3</td>
<td>12</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>vary</td>
<td>vary</td>
<td>22</td>
<td>3</td>
<td>9</td>
<td>9</td>
<td>4</td>
</tr>
</tbody>
</table>
### Readability

<table>
<thead>
<tr>
<th>Word</th>
<th>Attribu...</th>
<th>Total O...</th>
<th>Docum...</th>
<th>Wells F...</th>
<th>JPM...</th>
<th>Bank o...</th>
</tr>
</thead>
<tbody>
<tr>
<td>risk</td>
<td>risk</td>
<td>1097</td>
<td>3</td>
<td>421</td>
<td>363</td>
<td>313</td>
</tr>
<tr>
<td>exposure</td>
<td>exposure</td>
<td>291</td>
<td>3</td>
<td>82</td>
<td>84</td>
<td>125</td>
</tr>
<tr>
<td>approxi...</td>
<td>approxi...</td>
<td>119</td>
<td>3</td>
<td>40</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>exposur...</td>
<td>exposur...</td>
<td>97</td>
<td>3</td>
<td>20</td>
<td>33</td>
<td>44</td>
</tr>
<tr>
<td>risks</td>
<td>risks</td>
<td>92</td>
<td>3</td>
<td>38</td>
<td>31</td>
<td>23</td>
</tr>
<tr>
<td>assump...</td>
<td>assump...</td>
<td>94</td>
<td>3</td>
<td>40</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>unobser...</td>
<td>unobser...</td>
<td>54</td>
<td>3</td>
<td>18</td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>revised</td>
<td>revised</td>
<td>32</td>
<td>3</td>
<td>7</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>believes</td>
<td>believes</td>
<td>31</td>
<td>3</td>
<td>7</td>
<td>20</td>
<td>4</td>
</tr>
<tr>
<td>volatility</td>
<td>volatility</td>
<td>38</td>
<td>3</td>
<td>9</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>difference</td>
<td>difference</td>
<td>38</td>
<td>3</td>
<td>19</td>
<td>15</td>
<td>4</td>
</tr>
<tr>
<td>differ</td>
<td>differ</td>
<td>33</td>
<td>3</td>
<td>14</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>intangible</td>
<td>intangible</td>
<td>53</td>
<td>3</td>
<td>24</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>depend...</td>
<td>depend...</td>
<td>17</td>
<td>3</td>
<td>1</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>uncertai...</td>
<td>uncertai...</td>
<td>22</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>variable</td>
<td>variable</td>
<td>69</td>
<td>3</td>
<td>31</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>intangib...</td>
<td>intangib...</td>
<td>32</td>
<td>3</td>
<td>15</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>probable</td>
<td>probable</td>
<td>38</td>
<td>3</td>
<td>18</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Word</td>
<td>Attrubution</td>
<td>Total O...</td>
<td>Docum...</td>
<td>Wells F...</td>
<td>JPMorg...</td>
<td>Ban...</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td>----------</td>
<td>------------</td>
<td>-----------</td>
<td>--------</td>
</tr>
<tr>
<td>risk</td>
<td>risk</td>
<td>1097</td>
<td>3</td>
<td>421</td>
<td>363</td>
<td>313</td>
</tr>
<tr>
<td>exposure</td>
<td>exposure</td>
<td>291</td>
<td>3</td>
<td>82</td>
<td>84</td>
<td>125</td>
</tr>
<tr>
<td>exposur...</td>
<td>exposur...</td>
<td>97</td>
<td>3</td>
<td>20</td>
<td>33</td>
<td>44</td>
</tr>
<tr>
<td>approxi...</td>
<td>approxi...</td>
<td>119</td>
<td>3</td>
<td>40</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>variable</td>
<td>variable</td>
<td>69</td>
<td>3</td>
<td>31</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>assump...</td>
<td>assump...</td>
<td>94</td>
<td>3</td>
<td>40</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>conting...</td>
<td>conting...</td>
<td>29</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>24</td>
</tr>
<tr>
<td>risks</td>
<td>risks</td>
<td>92</td>
<td>3</td>
<td>38</td>
<td>31</td>
<td>23</td>
</tr>
<tr>
<td>intangible</td>
<td>intangible</td>
<td>53</td>
<td>3</td>
<td>24</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>believe</td>
<td>believe</td>
<td>41</td>
<td>3</td>
<td>25</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>unobser...</td>
<td>unobser...</td>
<td>54</td>
<td>3</td>
<td>18</td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>probable</td>
<td>probable</td>
<td>38</td>
<td>3</td>
<td>18</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>volatility</td>
<td>volatility</td>
<td>38</td>
<td>3</td>
<td>9</td>
<td>18</td>
<td>11</td>
</tr>
<tr>
<td>volatilities</td>
<td>volatilities</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>intangib...</td>
<td>intangib...</td>
<td>32</td>
<td>3</td>
<td>15</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>depend...</td>
<td>depend...</td>
<td>16</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>differ</td>
<td>differ</td>
<td>33</td>
<td>3</td>
<td>14</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>uncertai...</td>
<td>uncertai...</td>
<td>22</td>
<td>3</td>
<td>5</td>
<td>10</td>
<td>7</td>
</tr>
</tbody>
</table>
Polarity

UNIVERSAL STATES
SECURITIES AND EXCHANGE COMMISSION
Washington, D.C. 20549

FORM 10-Q

QUARTERLY REPORT PURSUANT TO SECTION 13 OR 15
(d) OF
THE SECURITIES EXCHANGE ACT OF 1934
For the quarterly period ended September 30, 2017

Commission file number 001-2979

WELLS FARGO & COMPANY
(Exact name of registrant as specified in its charter)

Motivation

Definition

Uses

Key Terms

Methods

Practice
Summary

- Textual analysis is useful in accounting today
- Basic textual analysis is done by computers – accountants' job is to be able to interrupt the results
- Students will know when it is appropriate to use textual analysis and what questions to ask when evaluating textual analysis results
Questions?
References


