Journal of Information Systems Education

Volume 35
Issue 1
Winter 2024

Teaching Tip

Using Text Analytics AI Insights in Microsoft Power BI Desktop to Score Sentiments, Extract Key Phrases, and Discover Unstructured Data Patterns

Chang Liu and Charles Downing

Recommended Citation: Liu, C., & Downing, C. (2024). Teaching Tip: Using Text Analytics AI Insights in Microsoft Power BI Desktop to Score Sentiments, Extract Key Phrases, and Discover Unstructured Data Patterns. *Journal of Information Systems Education*, 35(1), 48-55. https://doi.org/10.62273/PKER1800

Article Link: https://jise.org/Volume35/n1/JISE2024v35n1pp48-55.html

Received:	November 7, 2022
First Decision:	December 24, 2022
Accepted:	January 20, 2023
Published:	March 15, 2024

Find archived papers, submission instructions, terms of use, and much more at the JISE website: https://jise.org

ISSN: 2574-3872 (Online) 1055-3096 (Print)

Journal of Information Systems Education, 35(1), 48-55, Winter 2024
https://doi.org/10.62273/PKER1800

Teaching Tip
Using Text Analytics AI Insights in Microsoft Power BI Desktop to Score Sentiments, Extract Key Phrases, and Discover Unstructured Data Patterns

Chang Liu
Charles Downing
College of Business
Northern Illinois University
DeKalb, IL 60115, USA
cliu@niu.edu, cdowning@niu.edu

ABSTRACT

This teaching tip describes using Microsoft Power BI Desktop in a class to analyze unstructured data from an exit survey of prior students from a Master of Science in Management Information Systems program. Results from a short survey administered to these students showed that the students, using the no-code Power BI, were able to accomplish their text analytics tasks in a shorter period, and with less overall effort, as compared to using traditional code-rich text analytics methods. The process is described in detail as to how the students, new to unstructured data analysis, can uncover these results. This process can be replicated by other instructors teaching text analytics.

Keywords: Text analytics, Unstructured data, Artificial intelligence, Sentiments, Key phrases, Power BI desktop

1. INTRODUCTION

Questions such as *what are the major strengths and weaknesses of a degree program* and *how could a degree program be improved* are among many survey questions each graduate student is requested to fill out when graduating from a large Midwestern university. Those questions allow students to answer in an open-ended style that is not structured in a predefined data model. Their responses are stored in native text format as qualitative data. Those text or unstructured data present a challenge for businesses wishing to utilize their data for analysis (Balducci & Marinova, 2018; Gentzkow et al., 2019). Per a recent study from International Data Communications (IDC), 80% to 90% of organizations' data will be unstructured by 2025 (Rydning, 2021). In the meantime, an overwhelming majority of businesses (90%) say analyzed data is essential to their organization's innovation strategy (Krammer, 2021). Businesses can take advantage of text analytics tools, leveraging artificial intelligence (AI) including machining learning and natural language processing, to find meanings from enormous amount of unstructured data. Enabling smarter and scalable AI was identified as the top data and analytics trend by Gartner research (Panetta, 2021).

At the same time, businesses need to solve their data analytics problems quickly and cost effectively, and the no-code method has been shown to support accomplishing this goal by allowing them to promptly collect data, perform data analysis, and make accurate data predictions (Bhattacharyya & Kumar, 2021). Applying the no-code approach is transforming the traditional information system development paradigm (Wang & Wang, 2021; Wang & Wang, 2022). It makes it possible for both large and small businesses to deploy AI without hiring an army of developers, programmers, and data scientists (Reilly, 2021). Some even have suggested that the no-code AI platforms not only enable users to implement and test their ideas but also become the only way to ensure AI technology is adopted in every use case (Ataee, 2022; Smith, 2022). Therefore, there exists a strong appeal to teach students how to use AI insights without writing code to extract meanings, themes, and patterns from volumes of unstructured text data.

2. THE TEACHING ENVIRONMENT

The lesson described in this paper was based on a graduate business intelligence (BI) applications and tools course taught in Fall 2022 at a large, Midwestern university. All thirty-one students enrolled in this course were MIS majors. The lead author of this paper was the course instructor. The course was designed to introduce popular data analytics software tools from different vendors. Microsoft Power BI is one of those tools used by students for analyzing enterprise data. This was the second time this course covered the text analytics topic with a learning objective of applying the no-code method of AI technology to analysis of unstructured text data. (The text analytics topic was first introduced to the MIS and Data Analytics students in the Spring 2022 semester.) Twenty-nine out of thirty-one students were familiar with database management concepts such as entity relationship modeling,

database normalization, and SQL commands. Students either had or were taking a database course concurrently with the BI applications and tools course. However, twenty-two of the thirty-one students lacked solid programming skills, which is not unusual for MIS students in a business school.

Students were asked to use AI Insights of Text Analytics in Microsoft Power BI Desktop to score sentiments, extract key phrases, and discover data patterns from past MIS student exit surveys to present their findings on how the MIS degree program could be improved. Microsoft Power BI Desktop is an enterprise grade business intelligence (BI) software that lets users combine and analyze data from multiple sources to create interactive dashboards, reports, and apps (Ehrenmueller-Jensen, 2020). It is a popular data analytics tool used in academia (Techcronus, 2021). Although students can download Power BI Desktop for free, a subscription fee was paid by the school for the Power BI Premium Capacity so that students could use the AI Insights to gain access to its Cognitive Services that contain pre-trained machine learning models developed by Microsoft's science team. The AI Insights in Microsoft Power BI Desktop uses a natural language processing (NLP) technique to determine whether the sentiment expressed in textual data is positive, negative, or neutral (Microsoft, 2021). Microsoft Power BI Desktop is a powerful tool to help academic programs monitor student sentiments in feedback through student surveys and understand student needs. By using the pre-trained models, students can skip the steps for training, validating, and testing their models and move easily to analysis of their unstructured text data.

3. THE TEACHING APPROACH

Figure 1 shows the student exit survey source data loaded into Power Query Editor in Power BI Desktop. The data were stored in Excel. The data were edited by the course instructor to ensure no personally identifiable information was available to the students. With Power Query Editor, students learned to transform the data into the shape they wanted, e.g., by removing columns not needed for the study.

All software, including Power BI Desktop, used in this course has been installed in a virtual cloud computing environment called *Citrix*, or *Anywhereapps* as shown in Figure 2. This virtual cloud environment provides the students with access to the required software from anywhere at any time. This environment benefits the students during the COVID-19 pandemic since this approach allows the students to analyze data at their own pace, and from any location. Moreover, the instructor did not need to ask students to download the software and answer questions on software configurations.

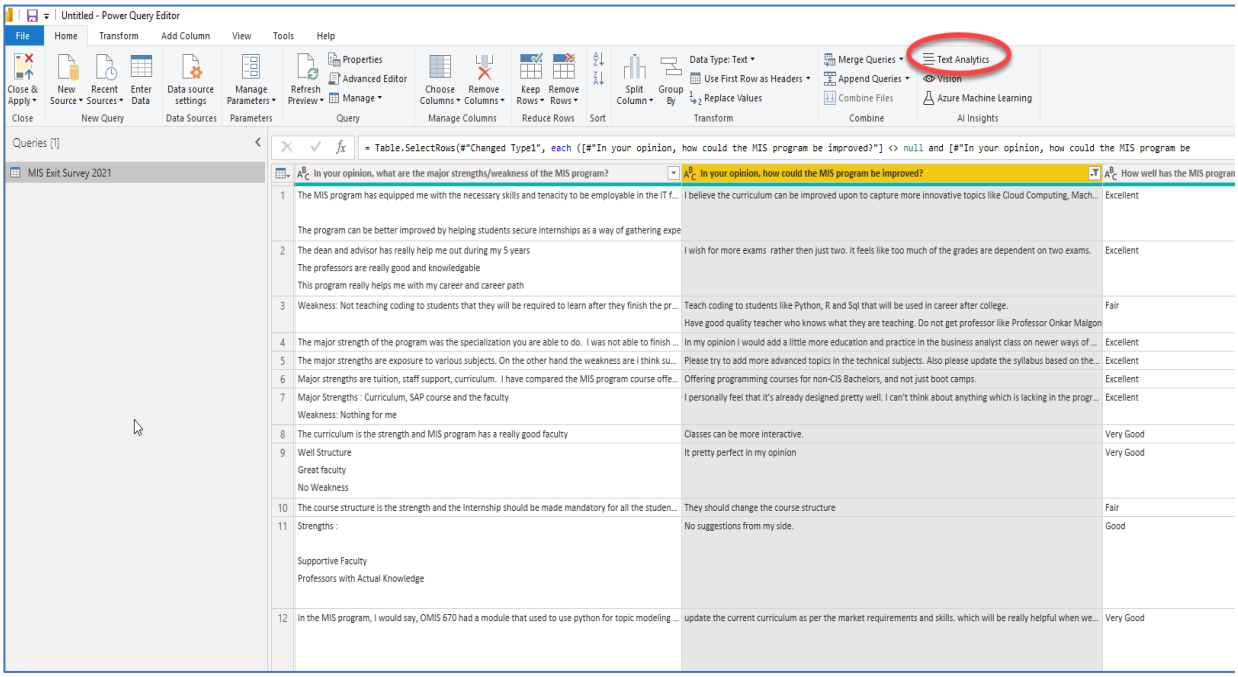


Figure 1. Student Exit Survey Data in Power BI's Power Query Editor

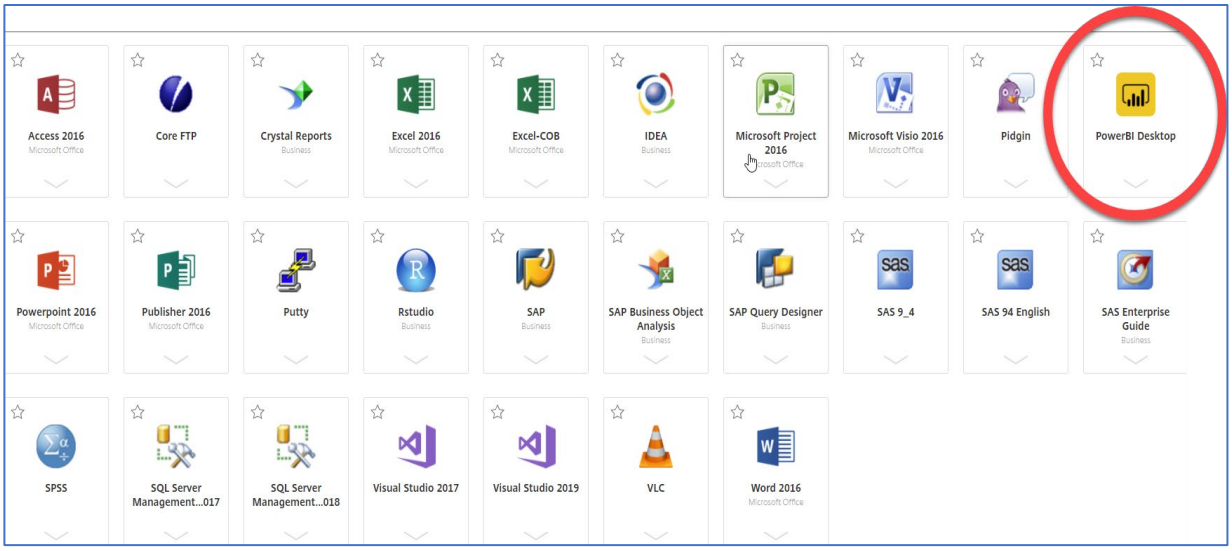


Figure 2. The Virtual Cloud Environment for Accessing Power BI Desktop

With the Text Analytics in Power BI, the students log into the Premium Capacity as shown in Figure 3. Then they can apply three different cognitive services, *i.e.*, Detect Language, Extract Key Phrases, and Score Sentiment as AI functions or insights shown in Figure 4. Figure 4 also shows that a premium license, *College of Business – OMIS*, was available for student use.

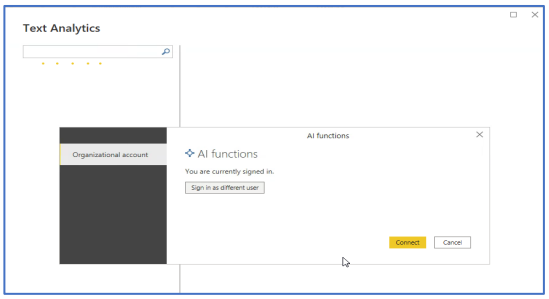


Figure 3. Login to Power BI Premium Capacity

3.1 Detect Language

The AI language detection function evaluates text input in the "*How could the MIS graduate program be improved?*" data column. After completing language detection, two new data columns, Language Name and Language Code Identifier (ISO), are created as shown in Figure 5. This function supports a wide range of national languages, variants, dialects, and some regional or cultural languages (Microsoft, 2021). The students then remove rows with *null* values based on rendered detected language name column. The instructor could have asked the students to delete those records with empty values for the analyzed text column in Excel since the original data were stored there. The deleting action after applying the AI function is a better way to help the students understand that the detected language serves as the input for other AI functions

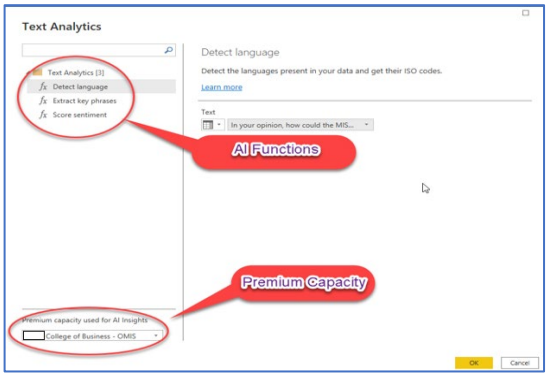


Figure 4: Three AI Functions for the Text Analytics in Power BI

3.2 Score Sentiment

Next, students continue to apply text analytics to score sentiment as described in Figure 4 for the question "*How could the MIS graduate program be improved?*". The AI engine's score sentiment evaluates the text input for the question and returns a sentiment score for each response, ranging from 0 to 1, *i.e.,* 0 represents extremely negative and 1 represents extremely positive. To be consistent with the school's program assessment criteria, the students are asked to create a column named "Sentiment Text" using the conditional formatting feature as shown in Figure 6 to specify the following outcomes: if sentiment scores ≥ 0.80, then "*Very Positive*"; if sentiment scores ≥0.60 but <0.80, then "*Positive*"; if sentiment scores ≤ 0.20, then "*Very Negative*"; if sentiment scores >0.20 but ≤0.40, then "*Negative*"; else the sentiment text will display "*Neutral*". Figure 7 displays the sentiment scores and the classified text values of "Sentiment Text" generated from this approach. The average sentiment was scored *0.68* or *positive* in this exercise.

Figure 5. Language Detection Results

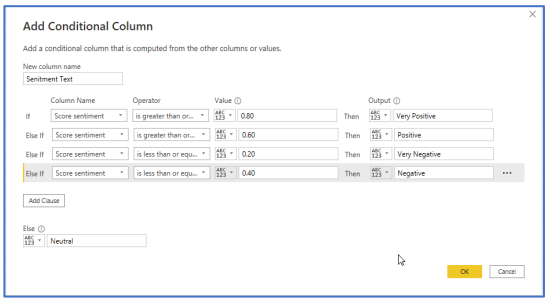


Figure 6. Group Sentiment Scores to *Very Positive*, *Positive*, *Neutral*, *Negative*, and *Very Negative*

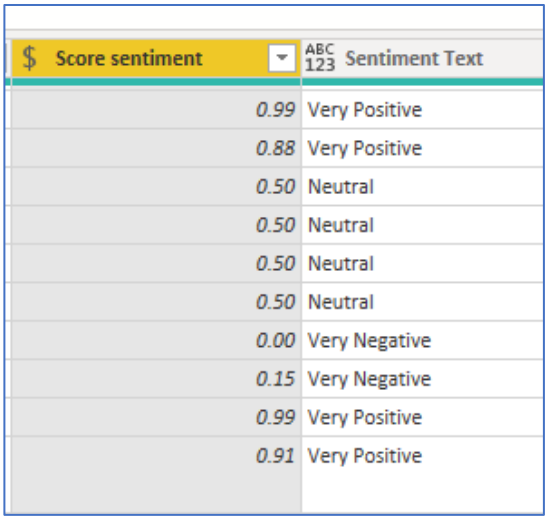


Figure 7. Score Sentiment and Corresponding Sentiment Text

3.3 Extract Key Phrases

Next, the students extract key phrases as described in Figure 4 for the question "*How could the MIS graduate program be improved?"*. The AI's key phrases extraction evaluates the unstructured text and returns a list of extracted key phrases as shown in Figure 8.

At this point, the students were excited to see sentiments scored and key phrases extracted from the survey question with this simple approach without any code writing. It was interesting to note that sentiment was scored as 0.50 or neutral if "not sure", "n/a", or "none" were provided as the feedback or input for the question. The students learned that if feedback is mostly objective there was no extracted sentiment, resulting in a 0.50 score. The students examined the extracted key phrases and they concluded that *Python programming*, *in-depth technical courses*, and *student-focused learning* emerged as the critical areas to be improved, based on the feedback from the exit survey.

These conclusions were drawn from a word cloud map students created which is presented in Figure 9. The word cloud visual has been used in various contexts to provide an overview of extracted text, indicating the frequency of occurrence of words (Heimerl et al., 2014). The more a specific word appears in a source of textual data, the bigger and bolder it appears in the word cloud.

3.4 Discover Data Patterns

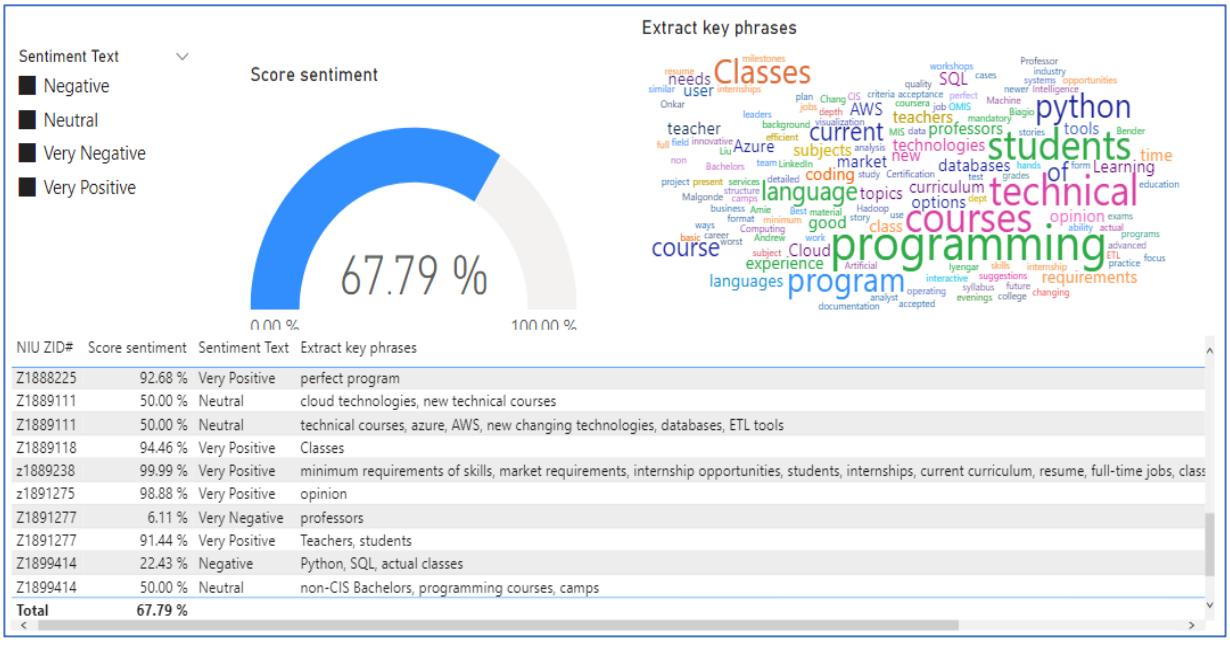
It would not be perfect to end the exercise by simply applying the AI insights to detect language, score sentiments and extract key phrases. The students were challenged further to present their text analytics results using various visualized reports. Figure 9 shows an example of a dashboard generated from the class exercise to show the average sentiment score in percentage, options (checkbox or slicer) to display the score sentiment for each sentiment text category, a word cloud map on the extracted key phrases, and a summary table of the sentiment score, sentiment text, and key phrases based on individual student id.

4. DISCUSSION AND SUGGESTION

An online survey was conducted by the authors (see Appendix) after completing the Text Analytics exercises. The students strongly agreed (Mean=4.81, Std. Dev=0.32, Median=5.0) with the statement that the *"no-code" approach of using Text Analytics of AI Insights accurately detected sentiments on "how could the MIS program be improved?"*. Furthermore, they felt more confident in the use of the "no-code" approach to solve Text Analytics problems (Mean=4.36, Std. Dev=0.68, Median=4.0).

Score sentiment	Sentiment Text	Extract key phrases
0.99	Very Positive	real time examples, skill, students
1.00	Very Positive	program, final project, Project management, python, SAS Enterprise M...
0.98	Very Positive	remedial Python course
0.99	Very Positive	examples, real time
0.88	Very Positive	classes, programming software, SAS, great programming, SQL, languag...
0.50	Neutral	additional stats course
0.50	Neutral	real job examples
0.00	Very Negative	data analytics topics, classes, major overhaul, major waste, tools nece...
0.15	Very Negative	subsequent courses, students similar, prepared students, computer sci...
0.99	Very Positive	people, hands data analytics, semester long capstone project, finish, jo...
0.91	Very Positive	advanced knowledge of data analytical applications, python
0.30	Negative	data-related applications, depth, python
0.89	Very Positive	program, courses, foundational class, capstone project, foundational p...
0.95	Very Positive	topics, depth, baseline, couple, knowledge
0.78	Positive	Better profs, longer timeline
0.50	Neutral	intro, Python

Figure 8. Extracted Key Phrases with Score Sentiment and Corresponding Sentiment Text



NIU ZID#	Score sentiment	Sentiment Text	Extract key phrases
Z1888225	92.68 %	Very Positive	perfect program
Z1889111	50.00 %	Neutral	cloud technologies, new technical courses
Z1889111	50.00 %	Neutral	technical courses, azure, AWS, new changing technologies, databases, ETL tools
Z1889118	94.46 %	Very Positive	Classes
z1889238	99.99 %	Very Positive	minimum requirements of skills, market requirements, internship opportunities, students, internships, current curriculum, resume, full-time jobs, class
z1891275	98.88 %	Very Positive	opinion
Z1891277	6.11 %	Very Negative	professors
Z1891277	91.44 %	Very Positive	Teachers, students
Z1899414	22.43 %	Negative	Python, SQL, actual classes
Z1899414	50.00 %	Neutral	non-CIS Bachelors, programming courses, camps
Total	**67.79 %**		

Figure 9. Discover Data Patterns From the Text Analytics

Teaching AI applications have become an integral part of machine learning and business intelligence courses in a MIS program, but many students and faculty have not completely explored it since familiarity with programming languages such as Python and R are often required. This is often particularly true in business schools where students would rather conduct data analysis without writing code. This paper describes a faster and an easier way to teach the AI insights of text analytics, compared to using a code writing approach, by using Microsoft Power BI to analyze unstructured text data. Interested readers can request the sample data file and lab instructions on the exercise from the lead author. Students showed great

enthusiasm to learn and apply AI technology using a no code approach. For example, immediately after the described exercise, the students were asked in class to apply the same AI functions, i.e., *Detect Language*, *Score Sentiment*, and *Extract Key Phrases* for the "What are the major strengths and weakness of the MIS graduate program?" survey question. This is another example of text input of unstructured data. Results indicated that all students could correctly analyze the responses without any additional help from the instructor. A homework assignment was also given in which students analyzed customer review comments from a local restaurant website. Again, students finished the assignment easily by using the approach described in this paper.

Moreover, the students were given extra credits to watch a YouTube video "*Simple Sentiment Analysis in Python*" to compare with the no-code method of Text Analytics in Power BI (see Appendix). All students either agreed or strongly agreed with the statement that the "no-code" approach of using Text Analytics in Power BI Desktop "solves the issue" faster and easier compared to code writing for conducting text analytics as demonstrated in the video.

The opportunity for analysts to give meaning to unstructured data is substantial. Although programming knowledge and skills are important, this no-code technique can help provide graduates to fill this labor need. As a result of the lessons learned, the authors would like to provide the following suggestions for other faculty to teach Text Analytics using Microsoft Power BI Desktop:

- Prerequisites for students should include the knowledge of database concepts for building table relationships. Students can learn to go deeper into the analysis if they have these skills. For example, the exercise could be extended to include additional files with student profile data after the personal identification information has been removed. Once the relationship is established using student campus id, sentiment scores could be compared based on major, gender, ethnic background, direct vs. transfer admission, or other dimensions. This additional demographic information will enrich program assessment for each academic program.
- Prior to applying text analytics in Power BI, it is important for students to cleanse and shape the source data to ensure data quality. Invalid or incorrect inputs should be filtered out for the data analysis. "*Null*" values should be removed after the language detection AI function is applied. Students can learn a valuable lesson on data quality if a faculty member guides students using Power Query Editor in Power BI to transform and cleanse the source data.
- It would be helpful to ask students to examine the aggregation or summarization method and data format for the sentiment scores. After the numbers were derived from "score sentiment," *Sum* was the default aggregation. Most students continued to use the default aggregation without realizing that using *Average* was appropriate instead. Students can enhance their learning if a faculty member guides them to review different types of summary statistics to aggregate data for analysis. Figure 10 shows the students were asked to use the *Average* and *Percentage* for the aggregated sentiment scores.
- Students should be challenged to further explore AI features in Power BI to include exercises on using *Q&A*, *Key Influencers*, *Decomposition Trees*, *Smart Narrative*, and *Vision Analytics* visual tools. Lab hands-on materials on these topics are available to share based on request. Along with text analytics, those exercises help students to gain a solid overview of AI features in Power BI.

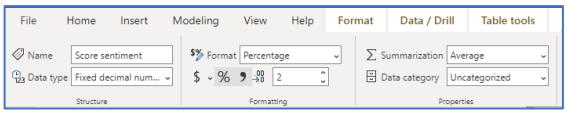


Figure 10. Change the Sentiment Score to Average and Percentage Aggregation

5. REFERENCES

Ataee, P. (2022). No-Code AI Platforms Bring AI to Everyone – Here Is How. *Medium*. https://towardsdatascience.com/no-code-ai-platforms-bring-ai-to-everyone-here-is-how-8f75b2f6ce9d

Balducci, B., & Marinova, D. (2018). Unstructured Data in Marketing. *Journal of the Academy of Marketing Science*, 46(4), 557-590.

Bhattacharyya, S., & Kumar, S. (2021). Study of Deployment of "Low Code No Code" Applications Toward Improving Digitization of Supply Chain Management. *Journal of Science and Technology Policy Management*, 14(2), 271-287. https://doi.org/10.1108/JSTPM-06-2021-0084

Ehrenmueller-Jensen, M. (2020). *Self-Service AI With Power BI Desktop – Machine Learning Insights for Business*. Apress.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as Data. *Journal of Economic Literature*, 57(3), 535-574. https://doi.org/10.1257/jel.20181020

Heimerl, F., Lohmann, S., Lange, S., & Ertl, T. (2014). Word Cloud Explorer: Text Analytics Based on Word Clouds. *2014 47th Hawaii International Conference on System Sciences* (1833-1842). https://doi.org/10.1109/HICSS.2014.231

Krammer, A. (2021). Embracing Data Analytics for More Strategic Value. *Harvard Business Review*. https://hbr.org/sponsored/2021/03/embracing-data-analytics-for-more-strategic-value

Microsoft (2021). *What Is Language Detection in Azure Cognitive Service for Language?* https://docs.microsoft.com/en-us/azure/cognitive-services/language-service/language-detection/overview

Panetta, K. (2021). *Gartner Top 10 Trends in Data and Analytics for 2021*. Gartner. https://www.gartner.com/smarterwithgartner/gartner-top-10-data-and-analytics-trends-for-2021

Reilly, J. (2021). How No-Code Platforms Can Bring AI to Small and Midsize Businesses. *Harvard Business Review*. https://hbr.org/2021/11/how-no-code-platforms-can-bring-ai-to-small-and-midsize-businesses

Rydning, J. (2021). *Worldwide Global DataSphere and Global StorageSphere Structured and Unstructured Data Forecast, 2021–2025*. IDC Research.

https://www.idc.com/research/viewtoc.jsp?containerId=US47998321

Smith, C. S. (2022). No-Code' Brings the Power of A.I. to the Masses. *The New York Times*. https://www.nytimes.com/2022/03/15/technology/ai-no-code.html

Techcronus. (2021). *Benefits of Power BI for the Education Industry*. Techronus. https://www.techcronus.com/blog/benefits-of-power-bi-for-the-education-industry

Wang, S., & Wang, H. (2021). Teaching Tip: A Teaching Module of No-Code Business App Development. *Journal of Information Systems Education*, 32(1), 1-8.

Wang, H., & Wang, S. (2022). Teaching Tip: Improving Student Performance by Introducing a No-Code Approach: A Course Unit of Decision Support Systems. *Journal of Information Systems Education*, 33(2), 127-134.

AUTHOR BIOGRAPHIES

Chang Liu is a Dean's Distinguished Professor of OM&IS in the College of Business at Northern Illinois University. Dr. Liu earned his Doctor of Business Administration in MIS from Mississippi State University in 1997. Dr. Liu's teaching and research interests are business intelligence, IT adoption and application, and data analytics. His research works published at *Information & Management*, *International Journal of Electronic Commerce and Business Media*, *Journal of Global Information Management*, *Journal of Internet Research*, *Journal of Computer Information Systems, American Journal of Business*, *International Journal of Mobile Communications*, *Journal of International Technology and Information Management*, and *Journal of Information Systems Education*.

Charles Downing is a Dean's Distinguished Professor and Distinguished Teaching Professor of OM&IS in the College of Business at Northern Illinois University. Dr. Downing researches, teaches and consults in Data Science and Big Data Analytics, Technology Strategy, and Information Systems Analysis, Design and Development. His articles have appeared in top journals, including *Communications of the ACM*, *Decision Sciences*, *Decision Support Systems*, *Information & Management*, *Journal of Global Information Management*, *Journal of Information Technology Management*, *Information Systems Management* and others.

APPENDIX

Text Analytics of Power BI Desktop Survey

Please rate your skills for the following:

	Poor	Fair	Average	Good	Excellent
(1) Programming – Writing Code	1	2	3	4	5
(2) Entity Relationship Modeling	1	2	3	4	5
(3) Database Table Normalization	1	2	3	4	5
(4) SQL Commands	1	2	3	4	5

Please watch the video of "Simple Sentiment Analysis in Python" at *https://www.youtube.com/watch?v=tXuvh5_Xyrw* (16 minutes)

Please complete the guided exercise of using Power BI Desktop for analyzing sentiments on "*How could the MIS program be improved?*" from past student exit survey.

Please share your opinion: **The "no code" approach of using Text Analytics of AI Insights accurately detects sentiments on "*How could the MIS program be improved?*"**

Strongly Disagree 1 2 3 4 5 Strongly Agree

Please share your opinion:

The "no code" approach of using Text Analytics in Power BI Desktop solves the issue in short period time with less effort compared to code writing for conducting text analytics as demonstrated in the video.

Strongly Disagree 1 2 3 4 5 Strongly Agree

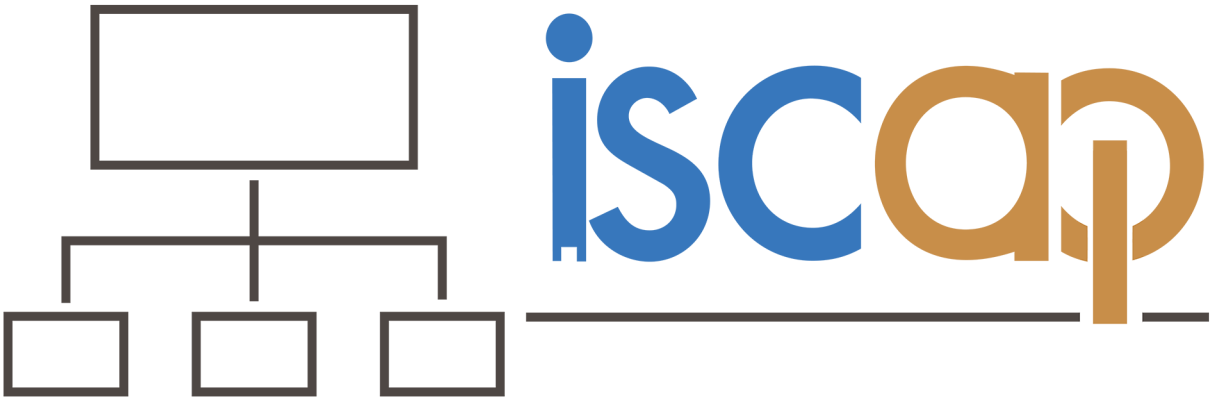
I prefer to use the "no code" approach to solve Text Analytics problems.

Strongly Disagree 1 2 3 4 5 Strongly Agree

I feel more confident to use the "no code" approach to solve Text Analytics problems.

Strongly Disagree 1 2 3 4 5 Strongly Agree

INFORMATION SYSTEMS & COMPUTING ACADEMIC PROFESSIONALS



STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the ***Journal of Information Systems Education*** have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN: 2574-3872 (Online) 1055-3096 (Print)