

Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Computing and
Information Systems

School of Computing and Information Systems

8-2019

EzLog: Data visualization for logistics

Aldy GUNAWAN

Singapore Management University, aldygunawan@smu.edu.sg

Benjamin GAN

Singapore Management University, benjamingan@smu.edu.sg

Jin An TAN

Singapore Management University, jinan.tan.2015@sis.smu.edu.sg

Sheena L.S.L VILLANUEVA

sheena.lv.2008@sis.smu.edu.sg

Timothy K.J. WEN

DHL

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research



Part of the [Artificial Intelligence and Robotics Commons](#), and the [Databases and Information Systems Commons](#)

Citation

GUNAWAN, Aldy; GAN, Benjamin; TAN, Jin An; VILLANUEVA, Sheena L.S.L; and WEN, Timothy K.J.. EzLog: Data visualization for logistics. (2019). *14th International Congress on Logistics and SCM Systems ICLC 2019, Taipei, Taiwan, August 19-22*. 1-8.

Available at: https://ink.library.smu.edu.sg/sis_research/4474

This Conference Paper is brought to you for free and open access by the School of Computing and Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Computing and Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

EzLog: Data Visualization for Logistics

Aldy Gunawan[†], Benjamin Gan, Tan Jin An, Sheena Lenore Sy Lee Villanueva

School of Information Systems

Singapore Management University, Singapore

Emails: aldygunawan@smu.edu.sg, benjamingan@smu.edu.sg, jatan@smu.edu.sg, sheenalv@smu.edu.sg

Timothy Kooi Jun Wen

DHL, Customer Solutions & Innovation – Asia Pacific

Email: TimotyKooi.JunWen@dhl.com

Abstract. With the increasing availability of data in the logistics industry due to the digitalization trend, interest and opportunities for leveraging analytics in supply chain management to make data-driven decisions is growing rapidly. In this paper, we introduce EzLog, an integrated visualization prototype platform for supply chain analytics. This web-based platform built by two undergraduate student teams for their capstone course can be used for data wrangling and rapid analysis of data from different business units of a major logistics company. Other functionalities of the system include standard processes to perform data analysis such as supervised extraction, transformation, loading (ETL), data type validation and mapping. Weather, real-time stock market and Twitter data can also be collected through EzLog’s web crawling function, as examples of external data that can be leveraged for more insights. Aiming to be user-centric, inputs from end-users were actively pursued in the design of the platform. Easily scalable, Logisticians can access the platform on their machines through Amazon Web Services (AWS) instances to perform descriptive and predictive analysis, including sentiment analysis and topic modeling, to better capture insights and identify patterns in logistics data.

Keywords: logistics, visualization, web-based application

1. INTRODUCTION

Logistics management is the process of strategically managing the procurement, movement and storage of materials, parts and finished inventory (and the related information flows) through the organization and its marketing channels in such a way that current and future profitability are maximized through the cost-effective fulfillment of orders (Christopher, 1998).

Recent studies in the field of big data analytics had resulted in the development of different tools and techniques to make data-driven logistics decisions (Govindan et al., 2018). By analyzing data and interpreting results in real time, decision-makers are able to make better and faster decisions to satisfy customer requirements.

Logistics which is a key enable of economies and global trade has been undergoing a vast degree of digitalization. This has provided an unprecedented amounts of data that can now be captured across supply chains to provide significant opportunities for optimizing capacity utilization, better serving customer needs, reducing risk, and even enabling business model innovations.

There have been several successful commercial logistics applications which help tackle problems such as network

optimization, transportation optimization and many more. However, many may not be suitable for users of smaller organizations because not all require a full-scale supply chain tool. As a result, use-case agnostic data wrangling tools such Microsoft Excel applications are commonly used in the industry which requires users to custom build solutions for their needs. Often, these tools have limitations in terms of data volume that it can handle and are unable to cope with the increasingly large data sizes available today. Furthermore, this approach is not easily scalable across multiple users, resulting in many variations of the same spreadsheet across teams in the case of using Microsoft Excel.

EzLog is a user-experience centric visualization platform that is designed to enable logistics companies to easily upload supply and demand information, clean and organize data, and perform data analysis. It aims to enable rapid extraction of business intelligence across very large datasets, as well as aggregate data from multiple sources including external data. The intent of the platform was also to be customized to the needs of the logistics industry specifically, and scalable to grow in analysis functionalities as requirements increase and become more sophisticated over time. The prototype developed by the student teams as described in this paper only includes several basic analysis capabilities to serve as a proof-

of-concept.

Many large Supply Chain tools focus on the functionality but neglect the design for usability. Lin (2009) studied a use case to implement a Supply Chain Management (SCM) system for a leading semiconductor manufacturing company in Taiwan. In the study, they discovered the need for usability and applied a usability framework to guide the implementation of the SCM systems. Although EzLog did not implement usability framework, it was developed with usability in mind throughout the design process. This was an essential design consideration as the platform needs to cater to a wide range of users with different technical competencies.

To achieve this, our students followed a design thinking approach to empathize with the users. They conducted several iterations using low and high-fidelity prototype to understand the user requirements. On top of that, they had conducted three rounds of User Testing (UT) to collect user feedback to further improve the system. Section 4 describes the UT and changes in scope based on user feedback.

Designing and redesigning the system based on the results of the UT was a challenging task as two different teams were working on the development of the same platform at the same time. During implementation, the student teams were required to design the component architecture to allow them to work separately on their codes and functionalities while sharing common modules such as account management, data pre-processing and upload. Although the plan was simple, execution was harder than expected. Some functions such as saving and exporting data as well as data filtering saw duplications across teams. The lessons learned from the student teams will be shared in the insights section.

The organization of the paper is as follow: Section 2 summarizes related works followed by a detailed description of the architecture and system design of EzLog in Section 3. Section 4 discusses the feedback obtained from the various UTs conducted with a logistics company and describes the various descriptive and predictive analyses. Finally, the paper is concluded in Section 5.

2. RELATED WORKS

The increasing amount of data available has motivated supply chain professionals to think about how data are produced, organized and analyzed. Hazen et al (2014) have highlighted the importance for organizations to adopt and perfect data analytic functions, such as data science, predictive analytics and big data, in order to enhance supply chain processes and performance.

Data visualization has played an important role in the advancement of modern data analytics. Bendoly (2016) presented frameworks to assist in the development of effective data visualization through contemporary theoretical notions of fit and sensemaking process. Pitfalls that can be caused by the

presence of bias are also considered and are exploited to ensure robust interpretability across stages and levels of use. The application of visualization systems faces issues in areas such as the transportation sector due to the lack of current tools that can efficiently handle tasks such as presenting abstract and detailed information, monitoring and analyzing order states. Hoen et al (2014) introduced a software system to visualize and demonstrate the model in the logistic of the transportation sector. It was also noted in the study that various domains of logistics and goods transportations utilize several specialized information management tools which leverage on available data and information for analysis. Additionally, Heydekorn et al. (2011) developed and implemented an interactive visualization in logistics for goods transportation to provide easy access to complex information spaces. The system included geo-referenced and time-referenced visualizations, which seamlessly consolidates diverse data sources and provides visual access to the heterogeneous and big data space. Furthermore, the system provides users with a user-friendly interface with search and filter functions.

Recently, research has been directed towards the use of social media for understanding its potential role in supply chain practice and research. Chae (2015) conducted a study on Twitter to develop an analytical framework to analyze *#supplychain* tweets and study the implications of using Twitter in supply chain contexts. Their research findings showed that supply chain-related tweets are more information-focused and conversational than regular tweets. This suggests that Twitter can be a medium for learning, promoting and networking. More specifically, a study into *#supplychain* tweets showed that Twitter is a widely used social media platform to improve stakeholder engagement, company promotion and hiring. Twitter data has also been used to identify supply chain management issues in food industries (Singh et al., 2018). More than one million tweets were collected and analyzed to identify customer satisfaction through sentiment mining. It is shown that the root causes of the identified issues were linked to their root causes in different segments of the supply chain. A set of recommendations for the development of a consumer-centric supply chain is then suggested.

Mass customization has become increasingly popular in many industries due to supply chain complexities and a high level of uncertainty and risks that companies face (Goh et al., 2013). The challenges in supply chain networks, being interconnected, non-linear and stochastic, have raised the need for an end-to-end supply chain visualization for efficient control and management of risks. For example, RiskVis is a supply chain visualization platform based on a multi-hierarchical modular design that displays spatio-temporal connectivity patterns of entities in a supply chain and accommodates real-time risk-related data collection and risk monitoring. RiskVis offers the flexibility to be customized

based on the user’s requirements – to process and store the supply chain data in the server and visualize the supply chain data.

EzLog combines data visualization, mass customization and the use of social media, weather and stock data in one platform. It assists logisticians in performing data analysis by providing a usable and interactive tool that performs supervised data cleaning, descriptive and predictive analysis, and data collection through web crawling. The next section describes the architecture and system design and detailed functionalities in the present platform.

3. ARCHITECTURE AND SYSTEM DESIGN

The overall architecture of EzLog is depicted in Figure 1. EzLog consists of four main components: data cleaning, crawling, back-end and front-end functionalities. Each component will be explained in the following subsections.

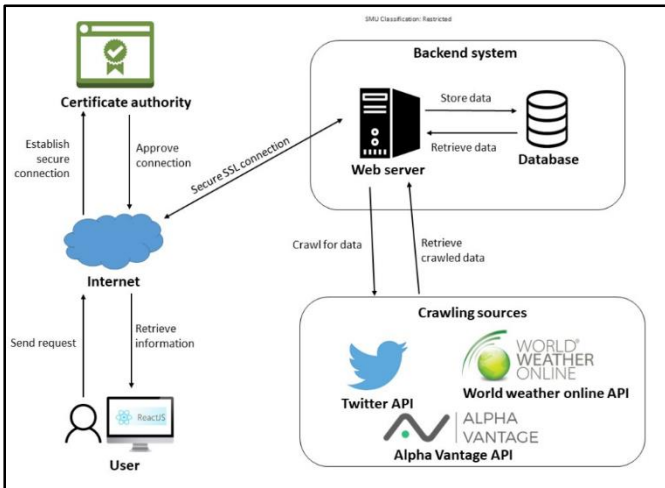


Figure 1: System architecture of EzLog

3.1 Data Cleaning

This component allows users to upload supply chain data into EzLog. It transforms and maps the uploaded data to make it more valuable for analysis. In the event of a misspelled header, it recommends a matching supply chain header using cosine similarity and Levenshtein Distance algorithm. Cosine similarity is used to calculate the distance between two words based on the magnitude of the vector. A drawback of the algorithm, however, is that it is unable to differentiate anagrams. To rectify this issue, the Levenshtein Distance algorithm was added as a secondary check to identify anagrams.

EzLog provides a list of predefined headers, that are commonly used in logistics for users to choose from. The tool provides a dropdown list of suggested headers, sorted from the

most similar to the least similar word based on the two algorithms. After validation, the data is then stored in a database hosted on Amazon AWS.

3.2 Crawling

EzLog allows users to crawl tweets, weather and stock information using open source public API by Twitter, World Weather Online and Alpha Vantage respectively. The data retrieved are correlated with logistics data for analysis. The following is a description of the user interface of the various crawling services in EzLog.

- **Twitter**
Users can enter any hashtag or keywords that they want, and enter the date to search. The system will then pull out a list of tweets based on the user's input from Twitter's API. The tweets can then be used for sentiment analysis.
- **Weather crawling (powered by World Weather Online)**
Users can specify the date range and country on the user interface to pull weather data out from the World Weather Online API. Because EzLog is an initial proof of concept software, it uses the free API key which limits the number of countries that can be selected.
- **Stock crawling (powered by Alpha Vantage API)**
Users are able to select a company from a list of predefined popular logistics company around the world. The system then uses Alpha Vantage API to crawl stocks information related to the selected company.

All crawled data are automatically saved in the database. In addition, the users are given an option of downloading the crawled data into their local machines as a CSV file.

3.3 Back-End

The back-end system is hosted on Amazon AWS to provide maximum uptime and scalability of the website. The algorithms used for data cleaning, statistical analysis, and data visualization are also hosted on the web server, using the Model-View-Controller framework to separate user interface from the backend handling.

Using React.js as client-side processing, the inputs made from the users are sent over to the backend system via an established SSL connection, certified by Let's Encrypt Authority, using "POST" method and received via Flask Python engine in the backend system.

3.4 Front-End

This component allows users to upload, share and analyze data. All users must have an account to access EzLog. Upon successful sign in, users will be directed to the home page of the system, as illustrated in Figure 2. Users are allowed to

upload one file at a time by choosing a file from the directory or to use the drag and drop function.

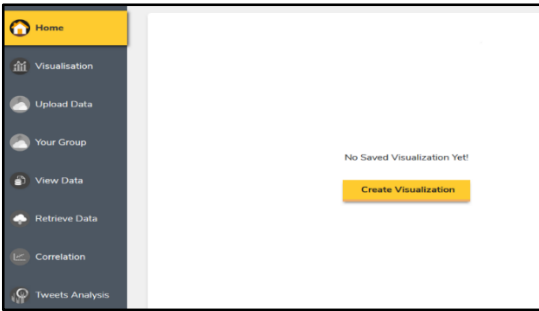


Figure 2: EzLog Homepage

EzLog has a capability to detect whether the uploaded file contains the correct headers. Correct headers are predefined by the company managers, based on a list of commonly used logistics headers. To maintain the quality of data in the database, the user can ask the group manager to add new headers into the approved list.

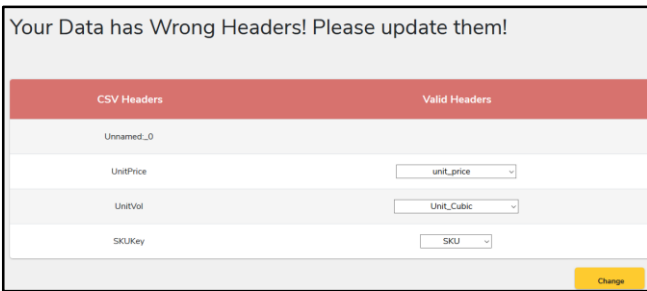


Figure 3: Illustration of wrong inputs

3.4.1 Descriptive Statistics

EzLog provides generated outputs for descriptive statistics. The system can display both bar chart and line chart to show descriptive values. Figure 4 illustrates an example of a bar chart that shows the top 5 SKUs for a set of customers.

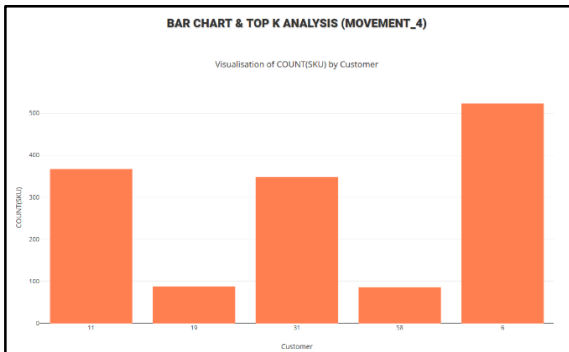


Figure 4: Bar chart of top 5 SKUs

3.4.2 Correlation Analysis

As an example of a type of analysis required in the Logistics industry, EzLog has the capability to calculate the correlation value between two different inputs. The correlation analysis is used to quantify the direction and strength of the linear association between two continuous variables. The correlation value ranges from -1 to 1. To determine the linear relationship between the dependent and independent variables, the regression equation is also included in the visualization. This can be used to predict the dependent variable for a possible value of independent variables.

Additionally, EzLog has a “Smart Correlation” function where it is able to generate different correlations from two different datasets. The correlation values are then sorted in descending order. Figure 5 provides a list of correlated values. We illustrate a high correlation value between Singapore weather and the quantity order of a particular SKU. The higher the mean temperature, the more quantity of the item is ordered.



Figure 5: Smart correlation output

3.4.3 Sentiment Analysis

EzLog also provides some inferential statistics to provide insights to the users. In recent years, Artificial Intelligence (AI) has become an important topic for analyzing big data. In EzLog, we have implemented one AI technique, known as Sentiment Analysis.

Sentiment Analysis is one of the text classification tools that automatically analyze the opinion about a given subject from written or spoken languages by categorizing the underlying sentiment as either positive, negative or neutral. Sentiment analysis is widely applied to various domains to capture reviews through social media pages or survey responses online.

Figure 6 illustrates an example of the sentiment analysis

for the word “logistics” from Twitter for a week. It shows that most of the tweets are neutral with a very low percentage of negative tweets. Those positive and negative words are listed in the word cloud as illustrated in Figure 7.

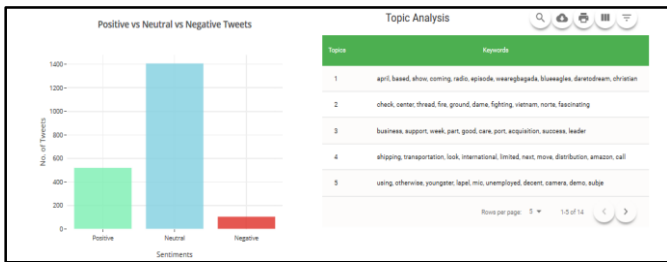


Figure 6: Sentiment Analysis output

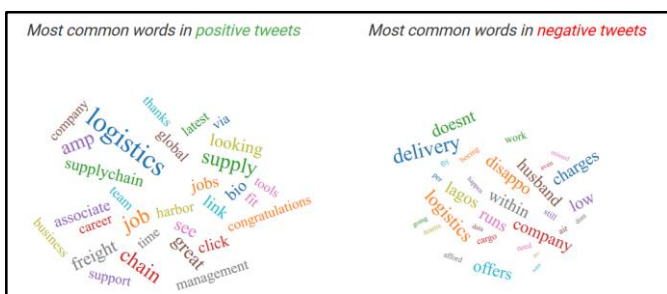


Figure 7: Word Cloud output

3.4.4 Topic Modelling

On top of that, we have included a type of statistical modeling called Topic Modelling to discover more insights about the crawled twitter data. Topic modeling is a frequently used text-mining tool for the discovery of hidden semantic structures in a text body.

Latent Dirichlet allocation (LDA) is a fully generative probabilistic topic model. This model plays an important role to capture latent topical information from a large collection of data (Ovsjanikov and Chen, 2010). The underlying idea of the probabilistic topic model is to cluster words that frequently occur together. By using contextual clues, topic models can connect words with similar meanings and distinguish between uses of words with multiple meanings.

An example of the topic modeling output is illustrated in Figure 6. Words are grouped into several topics, as shown on the right side of the figure. For example, "shipping-international-distribution" are grouped into one topic. It shows that most tweets use these words concurrently.

4. EXPERIMENTAL RESULTS

4.1 User Testing (UT)

In order to test the usability of the EzLog, students have

conducted three different User Testing (UT) sessions within three consecutive months. The purpose is to evaluate EzLog’s compliance with the project requirements and to verify if it has met the requirements of the business owner. Each UT comprises of a series of tasks and a feedback segment. In this section, we present the collected information during the UTs. A more in-depth analysis will be provided in Section 4.2.

During the first and second UTs (UT1 and UT2), users were asked to test the website and provide some feedback or comments. Most of the feedback received from UT1 were issues on the user interface (UI) as the users were unable to find the correct function to complete the task that was given to them. For example, users who are not trained in IT were confused with the term ‘web crawling’ and was unable to obtain historical weather data. Making use of the feedback received from UT1, the system was then enhanced with a clearer and simplified naming convention. Furthermore, hints were also added for user input fields and a search function was implemented on the ‘View Data’ function to enable quick searches on the data.

As a result, UT2 received better responses as compared to UT1 after the improvements on the UI were made. In UT2, additional functions such as Twitter crawling have been added to the system. The problems faced in this UT were mainly related to some further minor UI adjustments such as reducing the size of the table on the ‘View Data’ function. The average score (out of 10) for the ‘look and feel’ of the website given by the participants decreased from 2.17 to 1.40, with 10 as most difficult to use. Table 1 depicts the scores of the functions tested by the participants in UT1 and UT2 (1 as the easiest to perform the task for the function and 10 as the hardest to perform the task for the function).

| | UT1 | UT2 | Difference |
|------------------------------------|-----|-----|------------|
| View & combine datasets | 3.5 | 2 | 1.5 |
| Correlation | 2.6 | 2.1 | 0.5 |
| Web Crawling | 3.3 | 2.2 | 1.1 |

Table 1: Scores from UT1 and UT2

The final UT (UT3) was conducted with the following scenarios where each scenario was conducted by one group of students. Each user was then asked to evaluate the level of difficulty in a set of different tasks, as listed in Table 2.

Scenario 1: Keith monitors and reports the company’s inventory along with two others. Keith has excellent record keeping abilities with attention to detail, and a business mindset. He finds ways to optimize inventory control procedures and inspects levels of business supplies and raw materials to identify shortages. He ensures product stock is adequate and can cover customer demands. He plans the warehouse capacity and storage requirements of each product. When the stocks are low, he gets the logistics manager to place orders to replenish stocks. Sometimes in anticipation of

volume demand such as during festive seasons, he will replenish stock to avoid shortages. However, since he cannot predict the future, sometimes excessive surplus does fill up the warehouse. He decided to use the newly developed data visualization web application to monitor the movement of materials and whether there would be shortages or lack of space in the future. Furthermore, he uses the platform to work on the same dataset with his teammates.

to use the supply chain data visualization web application to see if the spike in goods movements and cost was related to the weather by doing some correlation analysis. Recently, you also realized that there have been some mentions on social media i.e. twitter about your company. In addition, you also noted that every time a new project is publicized on social media, the prices of the company's stock will change depending on the opinions of the public. Hence, you decided to retrieve the share price and tweets related to the company and perform sentiment analysis to have a sense of public sentiment towards the company and topic modeling to see what are the topics that the public are talking about.

Scenario 2: You are a manager from a supply chain department, and you want to better manage your resources for goods movement. Last year, you saw a huge increase in goods movement during the year end period (Oct-Dec) in Singapore, and coincidentally, it was also Singapore's hottest period in the past 10 years. The sudden huge increase in goods movement resulted in an increase in the cost of deliveries and also complaints from customers due to late deliveries as you had insufficient manpower to meet the demand. You want to avoid such issues in future, but before hiring more manpower, you decided

Table 2: Scenarios for UT3

| Scenario | Task | Description | |
|----------|------|--|--|
| 1 | 1a | You want to set up your team's dashboard in the platform. | |
| | 1b | You realize that you need to set the columns that are valid for each dataset. | |
| | 2a | You want to add the dataset given into the visualization platform in order to analyze the inventory data of the company so far. | |
| | 2b | You want to make sure that the headers and type of data in each column of the dataset you uploaded are correct. The correct headers of this dataset are the ones found in the dataset information. | |
| | 2c | After successfully uploading the dataset to your own dashboard, you would also like to share this with your whole team. | |
| | 3a | You want to quickly examine the top 3 product_names by quantity and unit_price without having to create your own charts or setting up any chart configurations. | |
| | 3b | You want to examine UAT_Data_Final.csv more carefully and see if there are any trends on shortages in inventory. Inventory shortages usually occur when there is less than 7 quantity for a Product_Name. | |
| | 3c | You realize that Product_Name Stone has been having shortages the past few days so you want to find out how it has been performing. Does it always have good supply levels? | |
| | 3d | You would now want to save this visualization so that you can view it in the future without having to repeat the procedures for creating it. | |
| | 4a | You would now want to figure out the top 3 Site_Names based on their quantity count. | |
| | 4b | You would now want to save this second visualization so that you can view it in the future without having to repeat the procedures for creating it. | |
| | 5a | Finally, you also want to save the dataset with the corrected headers so that you can use this data from now on and also share it with your team in the future. | |
| | 2 | 1a | Login using username: _____@gmail.com and password: 1234. You want to view the datasets content to answer some of the questions you have in mind. |
| | | 1b | You want to know the mean temperature for the particular movement with an order quantity of 108, to do that you decided to combine both datasets for easier visualization. |
| 1c | | You realized that the amount of Singapore's weather data in the database is insufficient. Hence, you decided to retrieve more Singapore's weather data (i.e. time period from "01/10/2018" to "31/12/2018") to do the correlation analysis. | |
| 1d | | You want to find out whether there's a correlation between the temperature in Singapore and the cost of movements for the month of December. | |
| 1e | | Instead of just finding out the correlation between the two variables (mean temperature and cost of movements), you decided to use the Smart Correlation function to find out whether there are any other variables that are highly correlated between the two datasets that you have just used. | |
| 2a | | You want to retrieve the share prices of the company to use them for the correlation analysis. | |
| | 2b | You want to retrieve 2000 tweets related to "company's name" for the past 7 days beginning from today's date to perform the sentiment analysis and topic modeling. | |

| | | |
|--|----|---|
| | 2c | You want to conduct a sentiment analysis based on the tweets that you've just retrieved to see the number of positive, neutral, and negative tweets. In addition, you want to find out what are the popular topics that the public is talking about related to the company. |
| | 2d | What are some of the most mentioned words from the positive tweets and the negative tweets respectively? |
| | 2e | You want to find out whether there's a correlation between the tweets' sentiment and the share price of the company. |
| | 3a | Explore the correlation function. |
| | 3b | Retrieve data from twitter using different tags and explore the tweets analysis function. |

4.2 Insights

In this section, we analyze the scores provided by the users during the UT3. The score for each task ranges from 1 (very easy) – 10 (very difficult). The average score for Scenario 1 is 4.03, whereas the average score for Scenario 2 is 2.19. The complete breakdown of the scores for each individual task in the two scenarios is illustrated in Figures 8 and 9. From the results, we observe that tasks in scenario 1 are more difficult to complete. This is expected as most of the tasks in Scenario 1 are related to data input and uploading. Most of the problems faced by the users are related to the upload data process as the limitation of free Amazon AWS server limits the speed of the server, resulting in a longer wait for the data upload process.

On the other hand, tasks in Scenario 2 appears to be smoother as it did not require large scale database uploading. The main functionalities tested in Scenario 2 are related to various statistical analysis such as correlation, sentiment analysis, and topic modeling. These functions rely on a mixture of client and server-side processing and did not have any major impact on the server's performance. Furthermore, Scenario 2 also tested on the crawling of data from external sources and it does not impact the performance of our own server.

Now, we further discuss the scores obtained for each scenario, as shown in Figure 8. In Scenario 1, tasks 1b, 2b, 2c, and 5a scored higher as they are related to the upload function. On the other hand, tasks 3d and 4b are scored lower as they are linked to the visualization function.

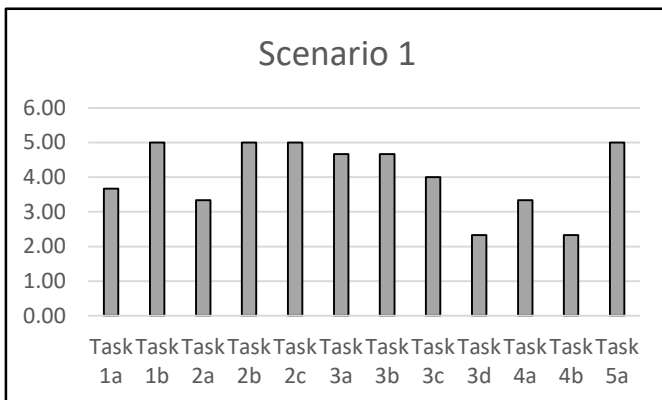


Figure 8: Scenario 1 scores

For Scenario 2 (Figure 9), on average, each task is scored below 3, except task 3b, which encourages the user to explore the crawling data function from the Twitter API. The main difficulty which is faced by the users is that they are unable to retrieve data from Twitter. This was an oversight from the student team as they are using the free Twitter API, which limits the number of records that could be retrieved from the twitter server. Furthermore, since the twitter crawling and sentimental analysis is a new function to the users, the users were very keen to crawl various data and they exceeded the free Twitter API limit within minutes of testing.

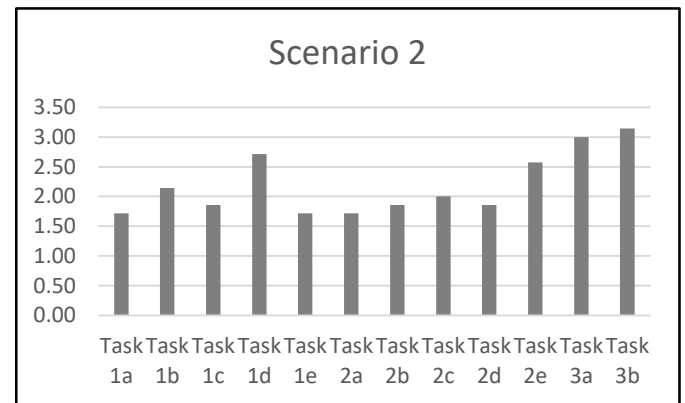


Figure 9: Scenario 2 scores

5. CONCLUSIONS

In this paper, we introduced EzLog, an integrated visualization platform for logistic needs. This platform was created and built by two undergraduate student teams for their capstone course. In this platform, users are able to upload, share the data and conduct some analyses. The user can perform the supervised extraction, transformation and loading (ETL) of data with data type validation and mapping of uploaded fields to logistic fields. In addition, EzLog provides web-crawling services to get online information, such as weather, Twitter and stock data. The component architecture of integrating Python code by reusing common modules and deploying to an Amazon Web Server (AWS) instances was described. EzLog also uses other tools and frameworks like React, Flask, Plotly, MySQL and various API interface.

In order to gather feedback to improve the performance of the platform, students have conducted three rounds of User Testing (UT). The latest version of EzLog will be demonstrated

during the ICLS 2019. Since this is the first version of the web-based platform, there is still room for improvement. With regards to future work, we plan to extend the performance of EzLog by adding more features for analyses and getting more insights from the data, such as multiple regressions and other statistical methods.

ACKNOWLEDGMENTS

This paper is based on a project from DHL-SMU Analytics Lab. We would like to thank the following team for their effort and contributions to the project: Andy ANG Yong Kiat, Bryan GWAN Wei Hao, Joel ANG Wee Keat, KIM Huiyeon, LE VAN Tuan Long, Patrick TAN Kun Wei, TAN Li Kai, Rainean Young CALUBAD, TAN Soon Wei, and XU Ying.

REFERENCES

- Hazen, B.T., Boone, C.A., Ezell, J.D., and Jones-Farmer, A. (2014) Data quality for data science, predictive analytics, and big data in supply chain. *International Journal of Production Economics*, **154**, 72-80.
- Chae, B. (2015) Insights from hashtag #supplychain and Twitter analytics: considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics*, **165**, 247-259.
- Singh, A., Shukla, N., and Mishra, N. (2018) Social media data analytics to improve supply chain management in food industries. *Transportation Research Part E: Logistics and Transportation Review*, **114**, 398-415.
- Govindan, K., Cheng, T.C.E, Mishra, N., and Shukla, N. (2018) Big data analytics and application for logistics and supply chain management. *Transportation Research Part E: Logistics and Transportation Review*, **114**, 343-349.
- Christopher, M. (1998) *Logistics and supply chain management: strategies for reducing cost and improving service*, New York: Prentice Hall.
- Hoen, P.J., Redekar, G., Robu, V., and Poutre, H.L. (2004) Simulation and visualization of a market-based model for logistics management in transportation. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, New York, USA*, 1218-1219.
- Heydekorn, J., Dachselt, R., Nitsche, M., and Nürnberger, A. (2011) On the interactive visualization of a logistics scenario: requirements and possible solutions. *Proceedings of International Workshop on Digital Engineering, Magdeburg, Germany*, 1-7.
- Bendoly, E. (2016). Fit, bias, and enacted sensemaking in data visualization: frameworks for continuous development in operations and supply chain management analytics. *Journal of Business Logistics*, **37** (1), 6-17.
- Goh, R.S.M., Wang, Z., Yin, X., Fu, X., Ponnambalam, L., Lu, S., and Li, X. (2013) RiskVis: supply chain visualization with risk management and real-time monitoring. *Proceedings of 2013 IEEE International Conference on Automation Science and Engineering (CASE 2013), Wisconsin, USA*, 207-212.
- Ovsjanikov, M, and Chen, Y. (2010) *Topic modeling for personalized recommendation of volatile items*. In Balc'azar, J.L., Bonchi, F., Gionis, A., and Sebag, M. (ed), *Machine learning and Knowledge Discovery in Databases, Lecture Notes in Computer Science (Springer)*, 483-498.