

The Application of Data Science on Food Waste Problem: A Preliminary Work

Amir Syazwan bin Ramlan, Iman Zulhakim bin Mohamad Dzain*, Raini Hassan

Department of Computer Science, Kulliyah of Information and Communication Technology (KICT),
International Islamic University Malaysia (IIUM), Gombak, Selangor, Malaysia.

*Corresponding author: imanzul0407@gmail.com

(Received: 1st November 2022; Accepted: 27th December 2022; Published on-line: 28th January 2023)

Abstract— Food waste refers to food that is fit for human consumption, but is discarded after it is kept past its expiration date or allowed to spoil due to factors like food spoilage, market surplus, and individual shopping and eating habits. It is a common problem worldwide, including in Malaysia. In reality, one-third of the total food generated is wasted or lost annually, and at least 820 million people worldwide do not have sufficient food. This study attempts to observe the current trends in the amount of food waste generated in Malaysia so that the public awareness can be increased. This paper also wishes to study about the undernourishment crisis for lowering the extreme hunger issues and to reduce the food losses especially rice, along the production and supply chain. This study indirectly aligns with sustainable development goals (SDG) related to zero hunger, and responsible consumption and production. Initially, the data science processes were implemented including data collection, data pre-processing, data cleaning, data exploration and selection, data wrangling, and exploratory data analysis (EDA). In Japan, the food waste issue was not given enough attention. It reflects the significant upswing in its early graph, which then began to decline. On the other hand, the food waste issue in Korea initially exhibited a sharp surge before a progressive drop. Regarding the undernourishment issue, almost all Southeast Asian nations nowadays are putting a lot of effort into eradicating or drastically lowering undernutrition among their citizens compared to the past. Rice is the common food type in Southeast Asia, which means it gets discarded frequently. The residents of Southeast Asia regard rice as a main food source.

Keywords— food waste, sustainable development goals, data science, exploratory data analysis, Southeast Asia, undernourishment, food type

I. INTRODUCTION

A. Background

To the best of knowledge, it is obvious that a portion of Malaysia's population, especially in the food service industry, lacks awareness on the opportunities and challenges of minimizing food waste, especially in the month of Ramadan. This is due to the poor management of some parties in handling food waste generation.

A Swiss startup has developed Sustainability-Oriented Innovations (SOI) technology, which was an ongoing development like Kitro for food waste management. This use of technology provides the services for hospitality, restaurants, and catering companies (HORECA) to address a new way to solve the food waste challenge. Besides, it also helps reduce wastage by collaborating with third-party companies to supplement the innovation activities [1].

However, some limitations were observed. A few business executives decided not to invest in technology after being interviewed. Apparently, some people were worried about being criticized because of possible bad waste measurements. Furthermore, small companies were

concerned about the incremental cost of invention. From the standpoint of society, they lacked proper knowledge of food waste management, but refused to take chances in exploring and practicing it as a part of their lives [1].

To relate this issue with Malaysia, an effort has been conducted by presenting the statistical data of current food waste trends that are getting worse, with the hope it may convince more businesses and companies to incorporate the innovation and increase public awareness regarding this crisis.

B. Research Objectives and Significances

The main goals of this paper are as follows:

- To provide insight into the current food waste situation in Malaysia.
- To observe the trends in the amount of food waste generated in selected Asian countries.
- To conduct in-depth research on undernourished people in Southeast Asia (except Singapore).
- To examine the top food types that are wasted by people in Malaysia, Indonesia, Thailand, and the Phillipines.

Significantly, the exploration of food waste using data science and machine learning can be considered an innovation. The study conveys the latest statistics and trends on food waste in Malaysia and its nearest countries through simple but visually appealing graphs. The findings will benefit society by raising awareness and concern about the current food waste generation, which is on the rise. It would be beneficial for the government to track the amount of food thrown away and provide some alternatives to society for reducing the risk of food waste.

II. RELATED WORKS

The number of studies on consumer-generated food waste has become more prevalent in developed economies. Liu [2] and her team evaluated mechanisms for Bangkok residents to prevent and reduce food waste phenomena. Several means were observing the latest food waste trend based on official published data, conducting a survey, and reviewing policies and strategies pertaining to food waste. However, the scenario of food waste over the past decade was still rising due to the incomplete law's stipulations. This can be seen through significant amount of food waste that more than doubled in 2018.

Mu'azu [3] found out that Saudi Arabia produced around 7.7 million tonnes of food waste in 2014. In fact, food waste generation was expected to strike accordingly as no concrete plans or food waste management-related policies have been devised. Aside from that, the major challenges were insufficient and challenging data collection, negative public attitudes, inadequate legislation, and lack of awareness about food waste management among citizens.

From the standpoint of operation management, an overview of current academic knowledge on the primary causes and remedies of food waste generation in merchants' supermarkets was presented. A management system based on the Plan-Do-Check-Act (PDCA) cycle was used to reduce the food waste level theoretically in identified food stores. Montama [4] claimed the fact that the study satisfies the requirements for a goal-oriented management system while preserving the PDCA cycle's features confirms its theoretical success. This management system seems to be suited for testing novel solutions that ought to significantly reduce food waste. Even so, the field had not seen much experimentation. By putting the suggested program into action, this still must be verified.

After COVID-19 struck, Heikal Ismail [5] and his fellow teammates conducted a descriptive statistical analysis to gain insight into the food waste trend before and during Movement Control Orders (MCO). Then, they proceeded with a one-way ANOVA test to examine the effect of MCO on food waste production in the Klang Valley area. They found that a 15% overall decrease in the amount of food

thrown away was shown during MCO. The ANOVA test also showed a considerable reduction that is less than 0.50. This infers that the food waste issue was less happening since people were prohibited from doing any outdoor activity during MCO took place, including dining in the restaurants, festivals, celebrations, buffets, and Ramadan bazars.

Looking at Malaysia's regional neighbour, Taiwan has enforced regulatory measures to make use of food waste to be included in the production of value-added resources, such as organic fertilisers and biogas-to-electricity. Tsai [6] observed the skyrocketing trend from 2003 until 2012. It started to decline slightly as the government highlighted the implementation of advertising strategies and regulations. Nonetheless, he also called for new regulatory measures and technological to be more environmentally friendly.

In another study, the random stratified sampling method and non-parametric t-test were implemented to elucidate the differences in food waste management in rural and urban Hanoi houses. According to reports, the food that was past its expiration date and declining quality led the residents in rural areas to waste more food than in urban areas. The study emphasized the importance of measurable and comprehensive policies targeting the entire supply chain, ongoing promotion to reduce food waste, and effective management systems [7].

The exploration of current food waste trends and situations creates an opportunity to conduct an in-depth study on the issue of undernutrition.

In relation to the food waste crisis, Mousa [8] had searched for and reviewed the alternatives to food security and food loss. She learned more about the malnutrition in Jordan that results from the problems of food insecurity (or lack of access to food) and food wastage. Referring to the latest report released by the Food and Agriculture Organization (FAO) on food security and nutrition, the rising rate of hunger worldwide is brought on by severe food insecurity, which affects the undernutrition crisis in the community [9].

On top of that, the prevalence of undernourishment (PoU) plays a crucial role in delivering the overview of this issue. It serves as a metric for the proportion of people whose food intake is consistently insufficient to meet their energy needs. Srinivasan [10] calculated the global PoU indicator from 1990 to 2017 with the implementation of the copula method to reconstruct the joint distribution of energy intake and energy requirements.

The food waste exploration can be extended to examine the issue of food types that are wasted frequently.

For instance, in a Polish supermarket, it was found that vegetables, fruits, meat, and fish products were the top contributors to the high mass of waste produced [11]. Whereas in Hungary, a study on quantifying the dominant

food types that were thrown away by households was heavily conducted. Soups and pottages were among the most frequently discarded food. The authors also stated that apple peel and non-consumable vegetable parts could become avoidable or unavoidable food waste [12].

In another research, Szakos [13] investigated the behavioural patterns of several consumers behind their household food waste with the execution of partial least square structural equation modelling (PLS-SEM). Throughout his findings, he observed that food types such as cooked meals, bakery products, dairy products, meat products, fruits, and vegetables were leading in the amount of food thrown away.

III. METHODOLOGY

The research was framed based on the data science lifecycle that includes conducting real-world data collection, processing data, cleaning data, wrangling data, implementing exploratory data analysis, modelling data, visualizing data, and building the data product. However, this study only focuses on up until exploratory data analysis.

A. Research Environment and Setup

- Python 3: The programming language Python 3 has been chosen as the main intermediary for the interaction with the data. Python 3 offers a wide range of libraries that can be used to ease the process of this research. For example, the Pandas library offers data analysis and manipulation tools; the NumPy library offers many features related to mathematical functions; and the Matplotlib library offers data visualization tools.
- Jupyter Notebook: The Jupyter notebook has been chosen as the main platform to accommodate the Python codes. It is an integrated development environment (IDE) that offers a web-based interactive development environment that allows the coding to be arranged in a flexible manner. It does support the Python 3 language, so it allows a project to advance progressively.

B. Data Collection

id	country	region	sp_code	commodity	year	loss_percentage	loss_percentage_original	loss_quantity	activity	food_waste_stage	treatment	cases_of_loss	sample_size	method_used	source	of_loss
0	MM	Myanmar	MM	1C	Groundnuts, excluding 209 shaded	5.22	5.22%	5.22%	(15)	NaN	Whole supply chain	NaN	NaN	FAO annual Agriculture Production Database	HC Source	NaN
1	MM	Myanmar	MM	1C	Groundnuts, excluding 209 shaded	5.43	5.43%	5.43%	(24)	NaN	Whole supply chain	NaN	NaN	FAO annual Agriculture Production Database	HC Source	NaN
2	MM	Myanmar	MM	1C	Groundnuts, excluding 207 shaded	5.61	5.61%	5.61%	(26)	NaN	Whole supply chain	NaN	NaN	FAO annual Agriculture Production Database	HC Source	NaN
3	MM	Myanmar	MM	1C	Groundnuts, excluding 208 shaded	5.4	5.40%	5.40%	(22)	NaN	Whole supply chain	NaN	NaN	FAO annual Agriculture Production Database	HC Source	NaN
4	MM	Myanmar	MM	1C	Groundnuts, excluding 205 shaded	5	5.06	5.06	(23)	NaN	Whole supply chain	NaN	NaN	FAO annual Agriculture Production Database	HC Source	NaN

Fig. 1 Food Loss and Waste Database dataset

There were three datasets that have been gathered. All datasets were collected from reliable sources, which are the

Food and Agriculture Organization of the United Nations (FAO) and the Organization for Economic Co-operation and Development (OECD). This research was focusing on the food waste situation in Malaysia. Thus, it is important to explore datasets that relate to Malaysia. Moreover, the data collection process includes data from Malaysia's neighbouring countries in the Southeast Asia region and even the Asian continent because of demographic and cultural traits. However, it is important to note that different countries might have factors that differ from Malaysia's food waste situation.

1) Food Loss and Waste Database by FAO

Referring to the Fig. 1, it can be observed that this dataset contains 3525 rows of data and 18 columns of features. The source database provides data from countries around the world, but instead it has been decided that only the Asia-Pacific region data be extracted as the research is focusing on Malaysia. This dataset can be considered a multivariate type. For each row of the data, it includes features like country, region, commodity, year, loss percentage, loss quantity, activity, food supply stage, cause of loss, sample size, method of data collection, and reference. The timeframe for every country is different, so the data might have been collected at different times.

2) Food Waste by OECD

This dataset contains 222 rows of data, and 17 columns of features as can be seen in Fig. 2. The source database provides detailed data for each country's food waste amount corresponding to a specific category. The data from Japan, Korea, and China were selected to focus more. The dataset includes features like country, category, activity, unit, source, year, and value (the food loss amount). The data were collected at distinct times since the timeframe might differ for each country.

LOCATION	Country	CAT	Category	ACT	Activity	UNIT	Unit	VAR	Variable	SRC	Source	TIME	Time	Value	Flag Codes	Flags
0	China (excl. Hong Kong SAR, Macao SAR)	FLW	1. Food loss waste	AGRICULTURE	Agriculture of products	PCT	%	PWRI15CN	Food Loss Rate (netton 00)	SOURCE34	Lunch/Villa & B	2010	2013	3.5	NaN	NaN
1	JPN	OTH	4. Others	FOODINDUS	Food Industry	TONNES	tonnes	PWRI15JP	Food Loss & Waste rate products unavailable for --	SOURCE37	Ministry of Agriculture, Forestry, and Fisheries	2011	2011	4621000	NaN	NaN
2	JPN	OTH	4. Others	FOODINDUS	Food Industry	TONNES	tonnes	PWRI15JP	Food Loss & Waste rate products unavailable for --	SOURCE37	Ministry of Agriculture, Forestry, and Fisheries	2012	2012	4818000	NaN	NaN
3	JPN	OTH	4. Others	FOODINDUS	Food Industry	TONNES	tonnes	PWRI15JP	Food Loss & Waste rate products unavailable for --	SOURCE37	Ministry of Agriculture, Forestry, and Fisheries	2013	2013	4870000	NaN	NaN

Fig. 2 Food Waste by OECD dataset

3) Prevalence of undernourishment (% of population) by FAO

In Fig. 3, it can be seen that this dataset contains 266 rows of data and 64 columns of features. This dataset provides data about the PoU, which is one of the indicators highlighted by FAO for the second SDG, zero hunger. This dataset includes features like country and the PoU value in

the form of a percentage of the population. The dataset has been retrieved and contains data on almost all countries worldwide.

Data Source	World Development Indicators	Unnamed: 2	Unnamed: 3	Unnamed: 4	Unnamed: 5	Unnamed: 6	Unnamed: 7	Unnamed: 8	Unnamed: 9	
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
1	Last Updated Date	08/04/2022	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
3	Country Name	Country Code	Indicator Name	Indicator Code	1960	1961	1962	1963	1964	1965
4	Aruba	ABW	Prevalence of undernourishment (% of population)	SM.ITK.DE.FC.ZS	NaN	NaN	NaN	NaN	NaN	NaN

Fig. 3 Prevalence of undernourishment dataset

C. Data Preprocessing

The dataset to use in the .csv file was imported using Pandas's `read_csv()` command and was stored in a dataframe. The number of columns and rows also can be obtained by using the `shape` command. In this paper, the missing values were checked first to successfully handle data from inaccurate inference and improper results. It was frequently applied to manage the null values before proceeding to the real data cleaning process.

D. Data Cleaning

The second data science process was data cleaning. Data cleaning involves the process of correcting irrelevant data and restructuring overall records to gain powerful insights as best as possible. It guarantees high-quality data that will produce an excellent result.

1) Food Loss and Waste Database by FAO

The data in several columns had null values, and they varied in number. This dataframe underwent the feature selection process first before it was decided if it needed further cleansing, as only a few columns will be included for the next data analysis.

2) Food Waste by OECD

There was no null value found in this dataframe. All columns were useful for later data exploration, except for the "Flag Codes" and "Flags" columns that will be filtered during the feature selection process.

3) Prevalence of undernourishment (% of population) by FAO

The dataframe left with exactly 266 rows of data after the correction of column header. Then the sum of each column's null values was checked with the command `isnull().sum()`. There were several columns that returned the sum of null values of 266, which means they were the empty columns. Hence, these columns were deleted. Moreover, the rows that had missing values were also filtered. The latest sum of

each column's null values was examined for a second time, and all features have no empty cells.

E. Feature Selection

Feature selection is the process of choosing the relevant columns that are sufficient to achieve the objectives or solve the problems. This is a better approach as only a small number of columns will be focused on during exploratory data analysis.

1) Food Loss and Waste Database by FAO

The dataset was used to achieve the third objective, which was to identify the top food types that were wasted among people in Malaysia and its nearest countries. The four columns from this dataset were very suitable for inspecting the data and its properties. The "country" column indicates the names of all territories involved, including Malaysia and other Southeast Asian countries. Furthermore, the "commodity" column represents the type of food waste produced. The "year" column denotes the observation of a period, and the "loss_percentage" column records the total amount of waste generated in percentage. With these features, the data can be visualized to figure out the food types that generated the highest amount of waste by their percentage losses over several years. On top of that, the resultant dataframe from this feature selection step was considered clean since neither column nor row had an empty value as in Fig. 4.

2) Food Waste by OECD

This dataset was taken to know the total amount of food waste generated in tonnes, which will be useful for descriptive analytics. There were duplicates of columns such as "Category" and "CAT", "Variable" and "VAR", and so on in the dataframe. Therefore, all duplicates, together with "Flag Codes" and "Flags" columns that have completely null values, were removed. The dataframe is now left with seven columns, refer to Fig. 5. The column header was reorganized to make it more understandable. From these records of data, it can be further visualized to identify significant patterns and trends. Examples of possible outcomes are knowing the total amount of food waste by different activities within a 12-year period, the trend of yearly cumulative waste and many more.

3) Prevalence of undernourishment (% of population) by FAO

The existing data was utilized to conduct an in-depth study of the undernutrition crisis, which aligns with the third objective. The focus of the study will be primarily on Southeast Asian countries, with the exception of Singapore, as the dataset does not include data from that country. After conducting several data cleaning processes, the paper

decided to drop several irrelevant columns, including “country code,” “indicator name,” and “indicator code,” as they did not contribute to the analysis of the trend of undernourishment. The dataframe now has 20 remaining columns, consisting of a column of country names and the 19-year observation for the rest, as shown in Fig. 6. These extracted features should provide actionable insights into malnutrition issues in terms of their severity for each year.

	country	commodity	year	loss_percentage
0	Myanmar	Groundnuts, excluding shelled	2009	5.22
1	Myanmar	Groundnuts, excluding shelled	2008	5.43
2	Myanmar	Groundnuts, excluding shelled	2007	5.61

Fig. 4 Food Loss and Waste Database dataset after feature selection

	Time	Value	Activity	Country	Category	Unit	Variable
0	2013	3.50E+00	Agricultural production	China (People's Republic of)	1. Food loss - Food waste	%	Food Loss from harvest (e.)
1	2001	4.64E+06	Food Industry	Japan	4. Others	tonnes	Food Loss & Waste and products unsuitable for ...
2	2002	4.83E+06	Food Industry	Japan	4. Others	tonnes	Food Loss & Waste and products unsuitable for ...

Fig. 5 Food Waste by OECD dataset after feature selection

Country Name	2001.00	2002.00	2003.00	2004.00	2005.00	2006.00	2007.00	2008.00	2009.00	...
0 Africa Eastern and Southern	33.648997	33.009286	32.602025	31.433222	30.209805	29.153501	28.791679	28.245725	26.912925	...
1 Afghanistan	47.8	45.6	40.6	38	36.1	33.3	29.8	26.5	24.4	...
2 Africa Western and Central	16.154919	15.717145	15.204831	14.54044	13.75271	13.171068	12.645863	12.3798	11.97082	...
3 Angola	67.5	83.2	58.7	55.1	52.2	49.3	48.1	42.9	33.2	...
4 Albania	4.9	6.1	7.5	8.6	8.9	8.8	8.3	7.3	5.9	...
...

Fig. 6 Prevalence of undernourishment dataset after feature selection

F. Data Wrangling

Data wrangling is the process of organizing unstructured data into a more valuable format for a variety of downstream purposes. This process ensures the data is of a high quality and can be effectively visualized and analyzed. It is an essential step in the data analysis process, as it allows users to transform raw data into useful statistics and insights

1) Food Loss and Waste Database by FAO

Since this study was mainly focusing on Malaysia and other Asian countries, the current dataframe had multiple countries from different continents as well. Thus, those non-Asian countries must be excluded first. In fact, Asian countries are also huge in number, so it is still left with multiple records and complex information to be further analyzed. Therefore, it has been decided that the taret countries are within the Southeast Asian continent. This was because Malaysia is located in that region, and the neighbouring countries practice quite similar lifestyles and food consumption patterns to Malaysia. A list of Southeast Asian countries has been established with the purpose of identifying the types of food that generate the highest amount of waste. The list was then used to filter the instances in the dataframe. The result was saved in a newly-created dataframe as it will significantly contribute to exploratory data analysis.

2) Food Waste by OECD

According to the overall data, there were multiple categories of food waste from different sources or activities, with their associated values and units. In addition, this information was recorded within the same years of observation between 2001 and 2012. From these points, perhaps a graph of the total amount of food waste produced broken down by activity can be envisioned. However, both households and post-production activities did not have any documentation about their associated amount of waste in 2012. This might be due to there being no data available in that year, or it was overlooked. A statistical analysis approach was implemented to cover this problem. A row was added to both dataframes of those two activities using the `loc()` method. The row was filled with the same items as the other rows in the dataframe. For the “Value” column, mean will be used to obtain a value accordingly. The mean is useful to come up with an expected value for both dataframes precisely and will not affect the overall results much.

3) Prevalence of undernourishment (% of population) by FAO

Several columns in the dataframe column header, particularly those representing years, were in float format. Therefore, a dictionary was created to replace the years in the float type with the integer type. The column header was renamed with the years that are in integer form. The scope of the malnutrition crisis was emphasized hence, the study was focused within the Southeast Asian continent. Another list of Southeast Asian countries was created for the purpose of examining the current trends of undernutrition issues. The list was then used to select the instances in the dataframe. The result was stored in a newly-created dataframe. For the purposes of proper analysis and

visualization, the new dataframe must be transposed. The country names, which were located on the left-hand side previously, are now placed in the first row. After that, the current column header was dropped and replaced with the values that were in the first row. The dataframe index was reset and the feature that represents years was renamed to "Year".

IV. ANALYSIS AND VISUALIZATION

The datasets that have been preprocessed will go through the EDA stage. This stage entails performing discovering patterns, checking assumptions, spotting anomalies, and developing hypotheses. Data visualization will be performed synchronously to assist the EDA process, using commands from the Matplotlib library in Python3.

1) Food Loss and Waste Database by FAO

In this dataset, the focal point to be analyzed is the top wasted food types within the Southeast Asia region. Unfortunately, it can only be deduced that the top wasted food type in Malaysia was rice for both 2007 and 2010, which recorded the same loss percentage as illustrated in Fig. 7. This situation infers that Malaysia does not have sufficient and reliable sources for data collection on food waste. Meanwhile in Indonesia, Fig. 8 depicted that vegetables and strawberries had the highest percentage loss among various food types within the years 2010 to 2018. The huge size of the country, as well as geographic and demographic differences from Malaysia's, were reflected in many food categories in its food waste list.

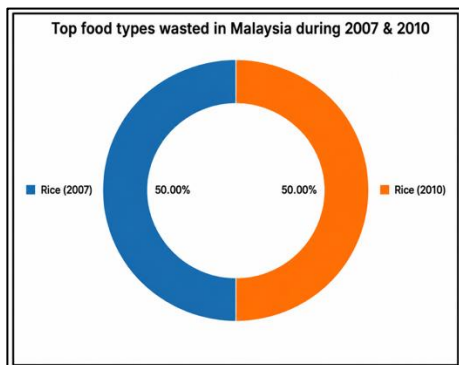


Fig. 7 A donut chart to depict top wasted food types in Malaysia during 2007 & 2010

On the other hand, rice during the year 2016 exhibited the highest loss percentage in Thailand, compared to cassava and rice from another year. Fig. 9 showed that the information was gathered between the years of 2013 and 2016. For the Philippines, it can be highlighted that maize was the food type that contributed to the sky-high loss of all food types in both 2011 and 2012. In Fig. 10, it noticeable that rice was behind maize as the second most discarded food.

Observing all four countries' top wasted food types with the measurement of loss percentage, they all have one food type in common, which is rice. To our best knowledge, rice is considered an important staple for Southeast Asian residents, especially in these four territories. Therefore, a multiple line graph in Fig. 11 was created to view and compare their trends with the rice loss percentage values.

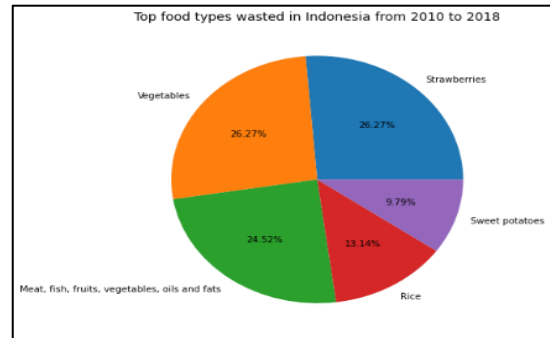


Fig. 8. A pie chart to depict top wasted food types in Indonesia from 2010 to 2018

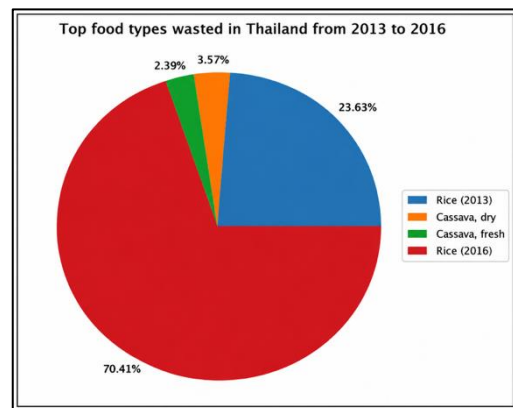


Fig. 9 A pie chart to depict top wasted food types in Thailand from 2013 to 2016

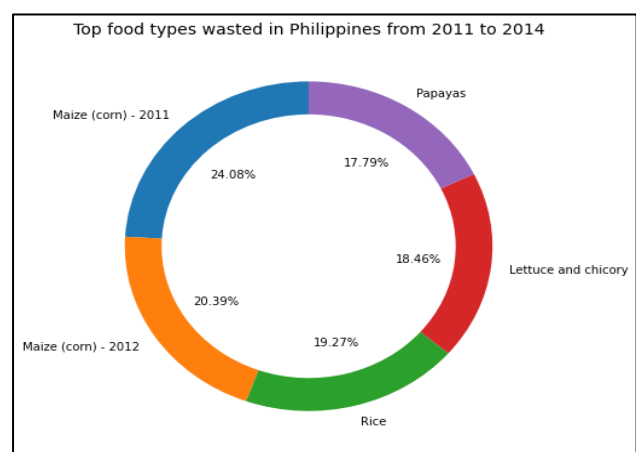


Fig.10 A donut chart to depict top wasted food types in the Philippines from 2011 to 2014

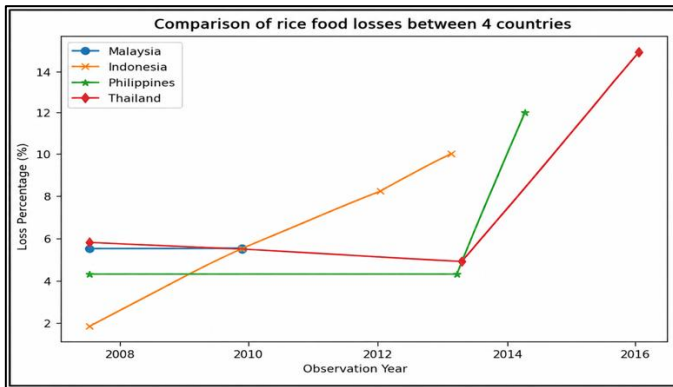


Fig. 8 A multiple line graph to depict comparison of rice that are wasted between 4 countries

All countries except Malaysia displayed an increasing trend by year, corresponding to their different observation times. Malaysia on the other hand, was constant and maintained the same loss percentage of rice. It can be assumed that these losses started even from the production stage in the paddy field until it served as food. There are many possible factors that resulted in the rice becoming spoiled, such as insects in paddy fields damaging the rice, mishandling in the packaging process, or inattentive food preparation.

2) Food Waste by OECD

Without the presence of Malaysia in this dataset, the analysis focus was on the total amount of food waste generated. Firstly, referring to Fig. 12, a multiple bar graph was generated to visualize the comparison of the total amount of food waste generated by selected activities from 2001 to 2012 in Japan.

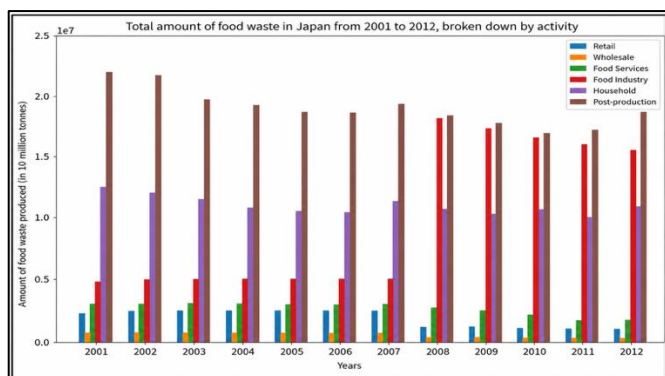


Fig. 9 A multiple bar graph to depict various activities that waste in Japan cause food from 2001 to 2012

All activities had ups and downs in the values, except a noticeably huge increase for the food industry from 2007 to 2008. Maintaining it at the same level is better than a huge jump, but this surely shows that the food waste situation in Japan was not treated as an important issue. Then, looking

at the line graph in Fig. 13 for the combined total amount of food waste generated by all activities. Initially, it decreased slowly from 2001 until 2006, then it began to skyrocket dramatically in 2008, before continuing a gradual decrease.

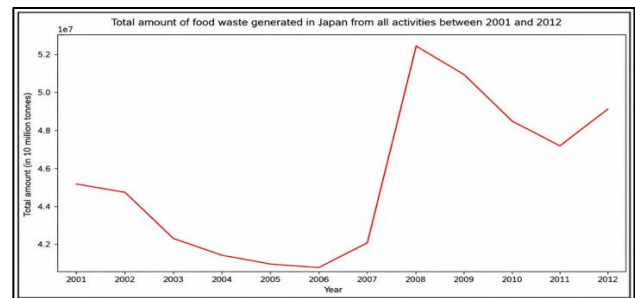


Fig. 10 A line graph to depict total amount of food waste generated from all activities between 2001 and 2012 in Japan

Next, the total amount of food waste generated yearly in Korea was also observed in a line graph in Fig. 14 depicted from 2004 until 2011. The total amount rose moderately from 2004 to 2008 before continuing to decline until 2011.

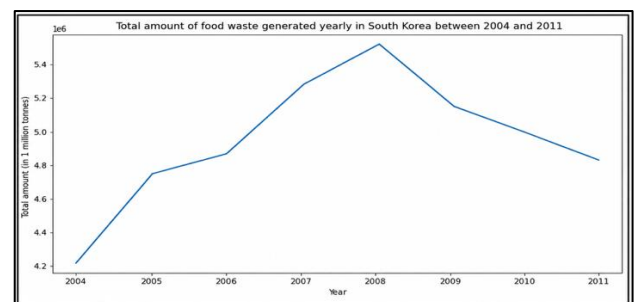


Fig. 11 A graph line to depict total amount of food waste generated yearly in Korea between 2004 and 2011

3) Prevalence of undernourishment (% of population) by FAO

The focal point that could be observed was the trend of undernutrition prevalence. A multi-line graph was depicted to show the trends of the percentage of undernourishment (PoU) for Malaysia and its nearest countries: Indonesia, Brunei Darussalam, Thailand, and the Philippines from 2001 to 2019. In Fig. 15, it can be deduced that all countries had a decreasing to steady trend. These trends indicated that those countries were focusing on eliminating or reducing undernutrition among their populations. Furthermore, a donut chart that depicts all ASEAN countries except Singapore has been established to gain proportional insights into the prevalence of undernutrition in comparison, as demonstrated in Fig. 16. It was established to get an overview of the undernourishment crisis in the Southeast Asia region, whose latest PoU was in 2019. All countries' PoU were noticeably under 10 percent of their country's

population, except for Timor-Leste, Thailand, and the Philippines.

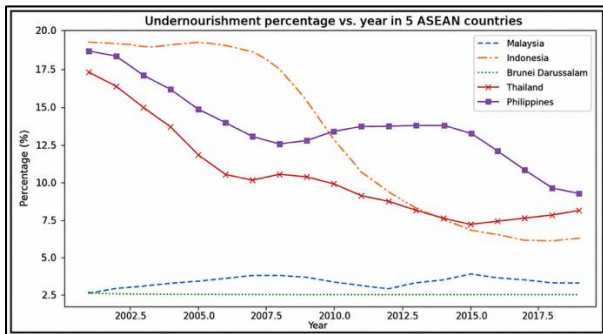


Fig. 12 A multiple line graph to depict undernourishment crisis between 5 countries from 2001 until 2019

Proportion of malnourished people in ASEAN countries in 2019

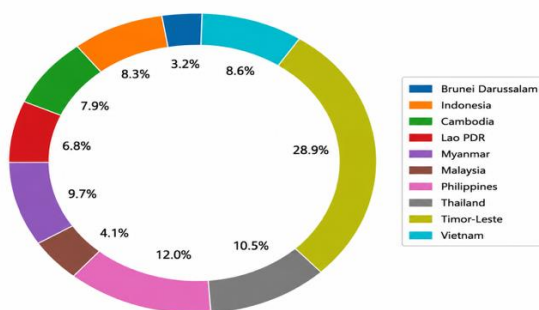


Fig. 13 A donut chart to depict proportion of malnourished people in ASEAN countries in 2019

V. CONCLUSIONS

From what has been observed in the EDA stage, it can be concluded that food waste has become a serious issue in the past and will continue to be if it is not treated. Food waste not only contributes to a huge portion of domestic waste, but it also causes a hole in food security for a huge percentage of the population. All objectives were achieved successfully in this paper. It has demonstrated the food waste trends in Malaysia and discussed the percentage of undernourishment and the frequently discarded types of food. The findings are useful, as the public will become more aware and concerned about the never-ending food waste and undernourishment crises. This paper also wishes to become the source of motivation for a lot more studies on food waste treatment and prevention. Throughout the study, some limitations were observed, including the lack of access to food waste data. It is very important to have reliable sources when finding suitable data for a great overview of the food waste issue. Even though some of the datasets that were used were imperfect, the data science processes that have been applied can transform the data

into an easily understood and interpretable format. For future work, the study hopes that the amount of food waste in Malaysia can be estimated or predicted with the help of machine learning algorithms since there has been no study that has carried out the experiment until now.

ACKNOWLEDGMENT

All praises are due to the Almighty. The authors are grateful to Kulliyah of Information and Communication Technology, International Islamic University Malaysia for their assistance and guidelines.

CONFLICT OF INTEREST

The author(s) declare that there is no conflict of Interest

REFERENCES

- [1] C. Martin-Rios, A. Hofmann, and N. Mackenzie, "Sustainability-oriented innovations in food waste management technology," *Sustainability*, vol. 13, no. 1, p. 210, 2020.
- [2] C. Liu, C. Mao, P. Bunditsakulchai, S. Sasaki, and Y. Hotta, "Food waste in Bangkok: Current situation, trends and key challenges," *Resour. Conserv. Recycl.*, vol. 157, p. 104779, 2020.
- [3] N. D. Mu'azu, N. I. Blaisi, A. A. Naji, I. M. Abdel-Magid, and A. AlQahtany, "Food waste management current practices and sustainable future approaches: A Saudi Arabian perspectives," *J. Mater. Cycles Waste Manag.*, vol. 21, no. 3, pp. 678–690, 2019.
- [4] C. Y. P. Montama, "Opening the food retailer black box, a proposal to reduce food waste in supermarkets." 2021.
- [5] M. Heikal Ismail *et al.*, "Impact of movement control order (Mco) due to coronavirus disease (covid-19) on food waste generation: A case study in Klang Valley, Malaysia," *Sustainability*, vol. 12, no. 21, p. 8848, 2020.
- [6] W.-T. Tsai, "Turning food waste into value-added resources: Current status and regulatory promotion in Taiwan," *Resources*, vol. 9, no. 5, p. 53, 2020.
- [7] C. Liu and T. T. Nguyen, "Evaluation of household food waste generation in Hanoi and policy implications towards SDGs target 12.3," *Sustainability*, vol. 12, no. 16, p. 6565, 2020.
- [8] T. Y. Mousa, "Food Security, Food Waste, and Food Donation and Redistribution in Jordan," *Austin J Nutr. Food Sci.*, vol. 8, no. 3, p. 1145, 2020.
- [9] M. Rosenberg, "Thoughts about food security, food loss and waste and what has to be done," *AIMS Agric. Food*, vol. 6, no. 3, pp. 797–798, 2021.
- [10] V. Srinivasan, "Prevalence of undernourishment in adults in 195 countries and territories." 2019.
- [11] B. Bilska, M. Piecek, and D. Kołozyn-Krajewska, "A multifaceted evaluation of food waste in a Polish supermarket—Case study," *Sustainability*, vol. 10, no. 9, p. 3175, 2018.
- [12] B. Szabó-Bódi, G. Kasza, and D. Szakos, "Assessment of household food waste in Hungary," *Br. Food J.*, 2018.
- [13] D. Szakos, B. Szabó-Bódi, and G. Kasza, "Consumer awareness campaign to reduce household food waste based on structural equation behavior modeling in Hungary," *Environ. Sci. Pollut. Res.*, vol. 28, no. 19, pp. 24580–24589, 2021.