The Philippines Under Fire: An Exploration of Fire Incident Open Data

Geremiah Marasigan
College of Computer Studies
De La Salle University
Manila, Philippines
geremiah_marasigan@dlsu.edu.ph

Ian Ona
College of Computer Studies
De La Salle University
Manila, Philippines
ian_ona@dlsu.edu.ph

***Abstract*—Fire Incidents are a big problem in the Philippines, and the country suffers greatly in damages each year because of it. In this research, we analyzed available open data on Fire Incidents, Firetrucks, and Internal Revenue Allotment in order to discuss methods to mitigate the damage caused by fire incidents in the Philippines. Through various data visualizations and statistical analysis, we found that more government funding could result in less damage from fire incidents while increasing the number of fire trucks may not aid in addressing fire incidents due to other external factors.**

I. Introduction

The Philippines has a history of disastrous fires, an example of which was the fire that broke out in the Kentex Manufacturing factory on May 13, 2015, that killed 74 people. Another major fire is the 1996 Ozone Disco Club fire that killed 162 people [1]. Disasters like these cost the country greatly in terms of infrastructural damage and lives lost. It is evident that proper disaster preparedness and fire incident handling are areas that the country can still improve on in order to safeguard the well-being of its citizens. Therefore, the Philippine government, specifically the Bureau of Fire Protection (BFP), should be better prepared so that it can respond quickly and efficiently to the next disaster.

Using open government data provided by the BFP regarding fire incident statistics, our study aims to provide a glimpse into the current state of the country in terms of fire incidents that occurred from 2012-2016. The exploration of these open data also aims to provide visualizations that will allow us to further understand possible relations between the number of fire incidents, injuries, deaths, and amount of estimated damages in Philippine pesos. Additionally, data on the Internal Revenue Allotment (IRA) dependency will be used to identify if there is a correlation between the amount of funding a local government unit receives from the government and their preparedness in dealing with disasters such as fire. The study will also explore the relation between the fire statistics and the current condition of BFP firetrucks in terms of their count, age, and cost. By the end of the study, we aim to have a deeper understanding on the possible factors that can be related to fire incident control as well as how the nation fares in dealing with fire disasters, as to provide insight into how we as a nation can better prepare ourselves for such disasters.

II. Literature Review

A. Fire Incident Analysis Internationally

The analysis of fire data is also an endeavor that other countries have already looked into. In 2015, a study [2] published in the Malaysian Journal of Forensic Sciences performed an analysis and visualization of fire incident data pertaining to Malaysia, Singapore, Indonesia, Great Britain, and the United States of America. Through said analysis, the study was able to compare the situation of fire breakouts in each of those countries. The study also compared how each of those countries were able to gather their data and how the different data gathering methods affected the completeness and depth of the data gathered. Finally, the study expressed the need for generalized fire statistics worldwide so that comparisons and assessments can be made to make sure that all beings on this planet are kept safe from the risk of fires.

In 2000, a study [3] was conducted to predict the outcome and effectiveness of fire brigades in London, England. The study used real data of current fire investigations as a basis. They examined the relationship of a fires threat and the presence of a firefighting force in the area. The study concluded that the presence of a firefighting force did not affect the probability of a fire occurrence in an area, however it did affect how the fire spread and therefore the amount of damage that the fire did. This study could be used to ascertain the effectiveness of firefighting operations as a way to help control the amount of damage that fires could do. We could then suggest the development of the firefighting forces in an area based on the data gathered as a possible method of limiting the damage that a fire occurrence could have.

A similar study [4] on determining the fire risk of areas through fire statistics was already done in China. In 2012, a study was published that analyzes the current fire statistics in China from 1991 to 2010. The study then uses this analysis to predict which areas have a higher probability of having a fire occur. They also analyzed the behavior of the fires and the damage they made to see how destructive a potential fire occurrence could have on certain areas. They were able to identify several factors that affected the risk of a fire occurring, including climate, economy, and the current fire safety management system in an area. They were also able to identify which factors made it more dangerous when a fire does occur, such as the density of an area, the time of day that the fire occurred, and even the time of year when the fire broke out. The study could be helpful as a reference or a guide on how to analyze fire data in the Philippines.

B. Fire Incidents in the Philippines

Compared to the previously mentioned countries, the Philippines is rather lacking in its efforts to minimize the damage caused by fire incidents. An article written by Mayuga [1] shows the current state of firefighting forces in the Philippines. According to the article, the country is lacking firefighters, due to the outdated equipment and the mistreatment of current firefighters. The Bureau of Fire Protection (BFP), the primary agency responsible for fire prevention and fighting fire, is short of fire stations, fire trucks, fire hoses, nozzles and breathing apparatus that could boost or enhance their firefighting capability. Additionally, BFP does not own the land under their fire stations, instead they have to rely on the generosity of Local Government Units (LGU) and other private entities to have land for their fire stations. The article shows the importance of finding ways to minimize the damage of fires in the Philippines. Our paper aims to help others become more aware of the current state of Filipino firefighters; how they can effectively help control fires and why its necessary for them to receive support from other Filipinos.

C. Related Models and Methodologies

In 2013, a comparative study [5] was done on two of the largest U.S. states by population (California & Texas) in order to identify which factors can affect the frequency and severity of fire incidents that occur. The study made use of fire incident data from the National Fire Incident Reporting System (NFIRS) that pertained to residential home fires that occurred from 2006-2010. The factors they considered to be possibly linked to fire incidents were areas of origin, item/s first ignited, and heat sources. The study made use of logistic regression modeling to compare the difference of the 3 factors between the 2 states and how it could affect the probability of a fire incident leading to a casualty. Our study can also make use of a similar methodology to compare certain factors between different regions or provinces that could possibly affect the number of fire incidents or casualties due to fire incidents in these areas.

The analysis of fire data is also recognized to be requisite in fire risk management and reduction. In 2010, Trabzon, Turkey was used as a pilot area for a study [6] that aimed to establish a fire incident database with the help of GIS or geographical information systems. The study recognized the need to aid local government units in decision making with regard to fire risk management in order to lessen or even eliminate the destructive effects of fire incidents. By having fire incident data available in a GIS-supported database, dynamic fire maps, visualizations, and queries could easily be made to help these LGUs with fire risk management. The system was also able to perform analysis on fire hydrant locations to analyze the needs of certain locations that lacked fire hydrants. Similarly, our study would also aim to incorporate geographical data to help analyze fire incident statistics across Philippine provinces, regions, or cities. Our study can also make use of similar analyses with separate datasets that pertain to the available firetrucks and stations of the Bureau of Fire Protection (BFP).

III. Methodology and Descriptive Statistics

For the study, we will be analyzing open government datasets from the Open Data Philippines web portal. The primary data set that we will be using is the Bureau of Fire Protections dataset for Nationwide Fire Incident Statistics from 2012-2016. The dataset describes the number of fire incidents per municipality, reported injuries, reported deaths, and estimated damages (in PHP) caused by fire incidents from 2012-2016. Only minor pre-processing will be done on the dataset such as proper title casing and spelling of cities in the Philippines. Other than that, it contains no null values, all columns are correctly typed, and column names are already descriptive. As a supplementary dataset, we will also be utilizing the BFPs dataset for fire truck information which details all the firetrucks of the country excluding the volunteer-owned fire trucks. We also plan to utilize the dataset for Internal Revenue Allotment (IRA) Dependency by province, provided by the Bureau of Local Government Finance, which details the share of revenues received by each province from the government.

For statistical testing, we will be performing inferential statistics and correlation testing that aim to answer the following questions:

1) Do areas with higher government funding deal with fires better and experience less estimated damages?
2) Does an increase in the number of fire trucks indicate an increase in their capacity to save lives and reduce damages? (Does there exist a correlation between the number of fire trucks and the deaths/estimated damages that occur? If so, what type of correlation?)

The study will also aim to answer the following questions using data visualizations:

1) How is the fire situation in the Philippines developing over the years?
2) Do fire-prone areas have more injuries and deaths than areas that are less fire-prone?
3) What could be possible factors that are related to fire incident count, injuries, deaths, and damages?

A. Descriptive Analysis

The dataset for Nationwide Fire Statistics from 2012-2016 has a total of 8169 rows x 10 columns, each entry describing fire statistics for a municipality in the Philippines for a particular year (see Table I).

TABLE I
DATABASE COLUMNS' DESCRIPTION

Column Name	Description
PSGC	The Philippine Standard Geographic Code (PSGC) is a classification and coding of the geographical-political subdivisions of the country, such as the region, the province, the municipality/city and the barangay.
Region	Region in the Philippines (1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12, 4A, 4B, ARMM, CAR, CARAGA, NCR)
Province Fire District	Districts as divided by the BFP
Congressional District	Legislative districts divided for representation in Congress
City Municipality	The city or the municipality of fire statistic
Year	Year of fire statistic (ranging from 2012-2016)
Incidents	Number of fire incidents
Injuries	Number of reported injuries sustained during fire incidents
Deaths	Number of reported fatalities
Estimated Damages	Estimated damages in Philippine Pesos caused by fire incidents

TABLE II

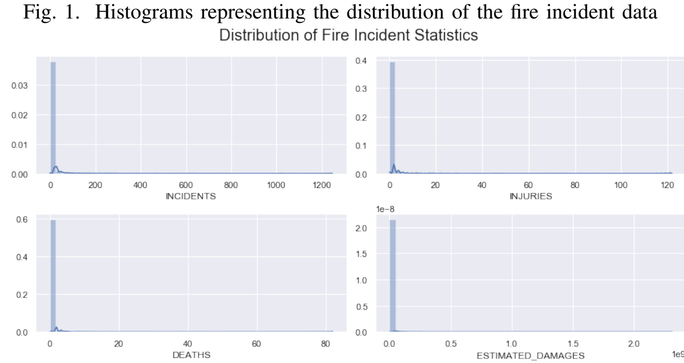
Column	Minimum	Maximum	Average
Incidents	0	1246	8.85
Injuries	0	122	0.50
Deaths	0	82	0.15
Damages (in millions of PhP)	0	2312.12	2.31
Fire Truck Acquisition Age (Years)	0	67	16.18
Fire Truck Acquisition Cost (in millions of PhP)	0	48.60	3.05

For the initial exploration of the data, we retrieved the minimum, maximum, and the average of the fire incident dataset and the fire truck dataset. From Table II, we can see that the maximum number of fire incidents is 1246 which is astonishing considering that the dataset's granularity was to the municipal level, meaning one municipality experienced 1246 fire incidents in one year. However, the mean count of incidents is 8.85, meaning that the max value was probably an outlier. Furthermore, the mean injuries and deaths are both less than one, indicating that the dataset contains a greater number of municipalities that experience 0 incidents in a year, leaving a smaller number of municipalities (presumably urbanized areas) as the main victims of fire incidents. The table also indicates that the average acquisition age of fire trucks is 16 years with an average cost of 3 million pesos, meaning that brand new firetrucks are not as commonly acquired by local governments.

Additionally, we will be observing possible relations between columns that we hypothesize to be related with each other. A scatter matrix will be used to observe possible relations between the following:

1) Incidents & Injuries
2) Incidents & Deaths
3) Incidents & Estimated Damages

Fig. 1. Histograms representing the distribution of the fire incident data



As part of the descriptive analysis, histograms were used to determine if the fire incident statistics followed a normal distribution. Figure 1 shows that for the four main fields of interest, they follow a right-skewed distribution, indicating that majority of the values lie to the right of the distribution's peak.

We will also be aggregating the dataset by province to find the average fire statistics per province over the course of 5 years. With this dataset, we will create a heat map visualization using geopandas so we may visualize which areas are prone to fire (higher number of fire incidents) as well as areas that have lost the most amount of money due to fire incidents (higher amount of estimated damages). Consequently, we hope to answer the following question/s with the help of these visualizations:

1) Is there a noticeable difference between the number of fire incidents in the different geographical island groups (Luzon, Visayas, Mindanao)?
2) Do more urbanized areas such as Metro Manila deal with fire incidents better than less urbanized areas?
3) Is an areas proximity to water a factor in determining the estimated damage a fire incident might have?

We will also be using a separate dataset from the BFP which details the distribution of fire trucks per region per province per municipality. Together with its corresponding visualization, we aim to identify if there is a relation between the number of fire trucks allocated to a province and the estimated damages that occur due to fire incidents in that province. We do this in order to answer the following questions:

1) How does the number of fire trucks allocated to a province affect that province's fire incident statistics (count, deaths, damages)?
2) How does average cost of fire trucks allocated to a province affect that province's fire incident statistics (count, deaths, damages)?
3) How does average acquisition age of fire trucks allocated to a province affect that province's fire incident statistics (count, deaths, damages)?

IV. Discussion and Results

A. Preliminary Analysis

Before going into geospatial data analysis, we first created some visualizations to answer the simple questions brought up in the previous section, as well as to provide a surface understanding of our data.

Fig. 2. A scatter matrix of the fields within the fire incident statistics

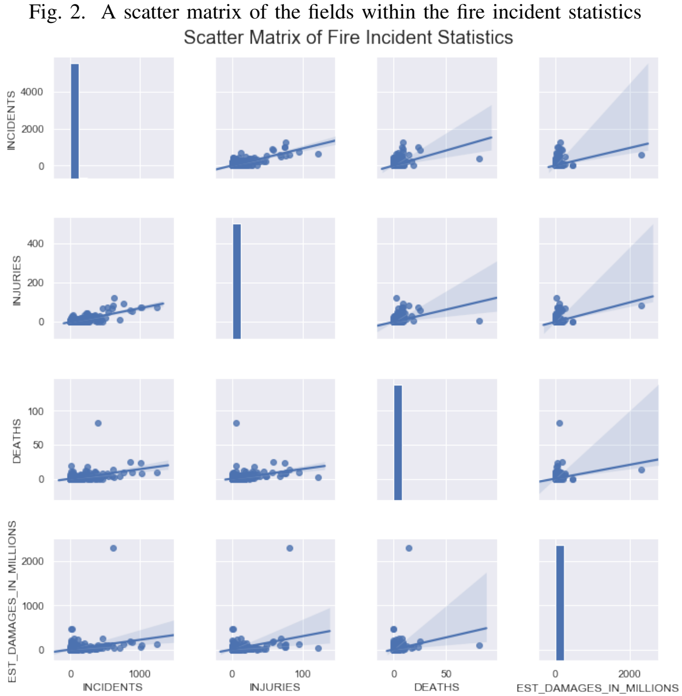
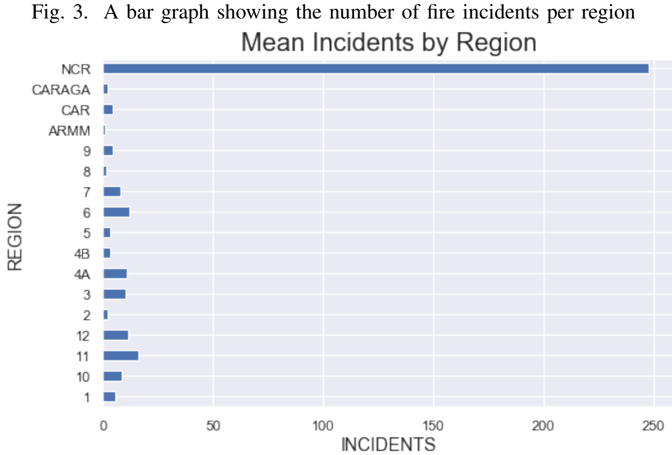


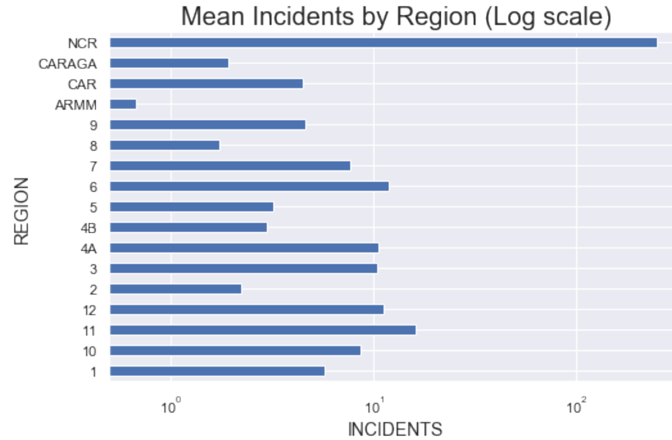
Figure 2 indicates that the fields within our primary dataset are all correlated to each other to varying degrees. This could be due to the interconnected nature of the fire incidents dataset in which a high number of fire incidents could lead to an increase in the number of fire injuries, deaths, and in turn, estimated damages.

Fig. 3. A bar graph showing the number of fire incidents per region



From Figure 3 we can see that by far most of the fire incidents occur in NCR, the most populous metropolitan area

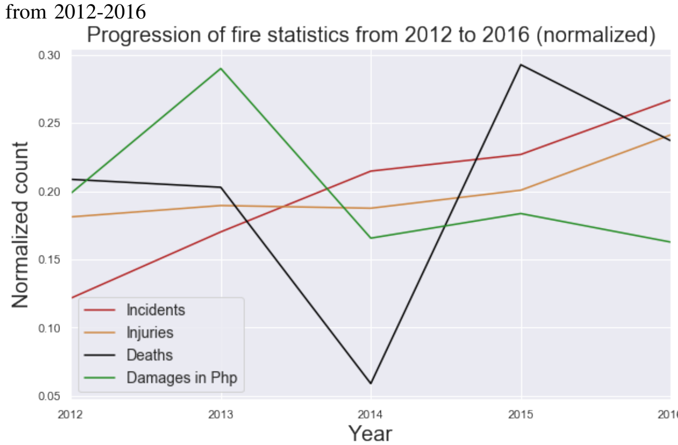
Fig. 4. A bar graph showing the number of fire incidents per region on a logarithmic scale



of the Philippines. This corroborates the studies done in the Literature Review. Since the Philippines' capital, Manila, can be found in this region, NCR is one of the most dense and congested areas in the country. As stated in Xin's study [4] and Mayaga's article [1], one of the leading factors to the damages caused by fires is the density of an area and the congestion on the roads. If the Philippine government wants Manila to be less fire prone, work could be done to manage the traffic in the area so that firetrucks would be able to easily move about in the city. The government could also work on developing the country's provinces so that it would be less densely populated in the cities.

To provide a better visualization of the mean count of incidents around the country, Figure 4 provides the same graph but uses a logarithmic scale. This is to properly show the differences in data when an extremely large value skews the scale of the visualizations. Through this visualization, we can see that next to NCR, Region 11 (Southeastern Mindanao) and Region 6 (Western Visayas) experience the most count of fire incidents.

Fig. 5. A normalized line graph representing the trend of fire incident statistics from 2012-2016

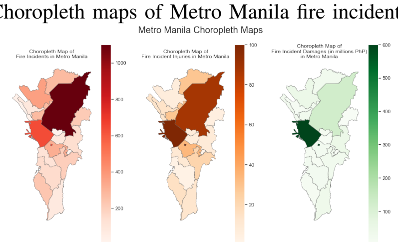


The last initial visualization done was a line graph that would help us observe possible trends in fire incident statistics throughout the years. It is evident that there is no clearly distinguishable trend in the number of deaths and the amount of estimated damages as the values for these fields rapidly increase and decrease with no identifiable pattern. However, we can observe a steady increase in the count of fire incidents and the number of injuries, which could indicate a slowly worsening situation of the country in regards to fire disaster handling and prevention.

B. Geospatial Analysis

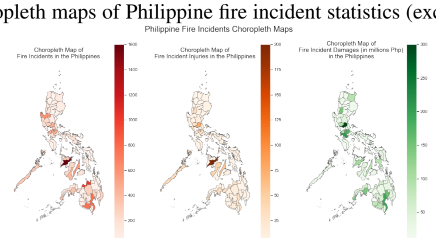
After doing the preliminary analysis, we could do some geospatial mapping to show if certain areas are more susceptible to fires than others.

Fig. 6. Choropleth maps of Metro Manila fire incident statistics



Due to the significantly higher amount of fire incidents in NCR, we have decided to create a separate choropleth visualization for Metro Manila as to not have the choropleth for the Philippines skewed to NCR. The maps in Figure 6 indicate that Quezon and Manila City experience the highest count of fire incidents. However, Manila City has slightly more injuries due to fire and also greatly surpasses Quezon in amount of estimated damages. This could be due to Manila being a more congested city with a population density more than twice that of Quezon City.

Fig. 7. Choropleth maps of Philippine fire incident statistics (excluding NCR)



From Figure 7, one could see that the areas with more incidents are also generally the same areas with more injuries and a higher amount of damage from fires. You could also see that there doesn't seem to be a relationship between a region's proximity to the ocean and the number of fire incidents that occur. An interesting finding is that even though Iloilo experiences the most number of fire incidents, areas in Luzon such as Bulacan, Batangas, Laguna, along with areas in Mindanao such as the Davao region, have a higher amount of estimated damages. This is possibly due to these areas having higher property value as they are moderately developed and urbanized areas that lie outside of the nation's capital region.

C. External Factors

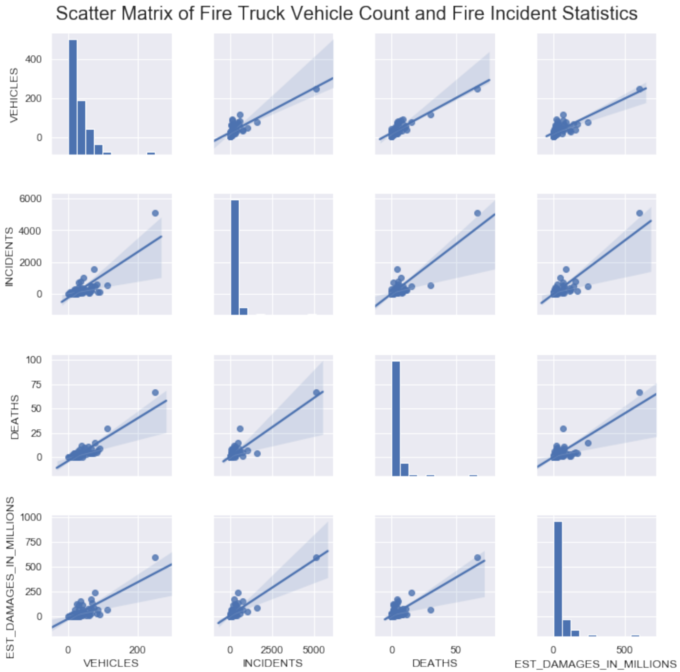
Other datasets were also analyzed to determine if there are other external factors that affect the damages or the number of fire incidents in the Philippines.

Fig. 8. Plotting of average IRA with various variables of the original BFP dataset

From Figure 8, we could see the relationship between the IRA (Internal Revenue Allotment) of a region and the effect of fire incidents on the region. It is seen that the scatter plot of the average IRA and the fire incident statistics has a best fit line that slopes slightly downwards, indicating an inversely proportional relationship between the local government budget and the effect of fire on that region. Since there also seems to be several points that diverge from the line, statistical testing will later be done to confirm the relationship between IRA and fire incident statistics. It is hypothesized, however, that the higher the IRA of the region or province, the more equipped they would be to handle fire incidents, indicating lower values for their fire incident statistics.

From Figure 9, we could see the relationship between the number of firetrucks in a region and the effect of fire incidents on the region. Surprisingly, we can see a direct relationship between the number of firetrucks and the effect of fire incidents in the region. This could be because of the previously stated hypothesis that metropolitan areas have more devastating fires, even though they have more firetrucks. Furthermore, an increase in the number of fire trucks does not necessarily indicate an increase in disaster preparedness as disasters may strike in areas located far from fire stations. These fire trucks may also be in varying conditions as the average age of these fire trucks during their acquisition is 16 years.

Fig. 9. Plot of the number of BFP firetrucks with various variables of the original BFP dataset



D. Statistical Testing

To determine the correlation between the variables we hypothesized to be correlated, statistical tests were conducted. Firstly, we performed Anderson-Darling tests to determine how the data is distributed. The null hypothesis states that: **Variable x follows a normal distribution,** and the significance level used was $\alpha = 5\%$. The results of each variable tested can be seen in Table III. With the Anderson-Darling test, we reject the null hypothesis if the resulting test statistic is higher than the critical value.

TABLE III
ANDERSON-DARLING TESTS

Variable Name	AD test statistic	Critical Value at 5%	Conclusion
Average IRA	9.31716	0.752	Reject the Null Hypothesis
Estimated Damages	2.2613	0.752	Reject the Null Hypothesis
Incidents	8.23156	0.752	Reject the Null Hypothesis
Vehicles	4.94807	0.752	Reject the Null Hypothesis
Deaths	7.89287	0.752	Reject the Null Hypothesis

Since Table III indicates that all the variables tested had their null hypothesis rejected, we can conclude that these variables do not follow a normal distribution. This is due to the fact that the data follows a discrete distribution. In natural disasters, most data follow a heavily tailed distribution [7]. Therefore, we will use Spearman's Rank Test to find the correlation of the data. The null hypothesis states: **There is no monotonic relationship between Variable x and Variable y,** and the significance level used was $\alpha = 5\%$. The results of various statistical tests for different pairs of variables can be seen in Table IV. With Spearman's Test, we reject if the p-value is lower than the alpha.

TABLE IV
SPEARMAN'S RANK TEST

Variable X	Variable Y	p-value	Rho value	Conclusion
Average IRA	Estimated Damages	2.50894e-3	-0.333421	Reject the Null Hypothesis
Average IRA	Incidents	2.15259e-4	-0.402415	Reject the Null Hypothesis
Vehicles	Deaths	2.20338e-17	0.771563	Reject the Null Hypothesis
Vehicles	Estimated Damages	4.76734e-17	0.766479	Reject the Null Hypothesis

A negative Rho value determines the direction of the correlation. The tests in Table IV show that there are statistical correlations between **Average IRA & Estimated Damages** (negative correlation), **Average IRA & Incidents** (negative correlation),**Vehicles & Deaths** (positive correlation), and **Vehicles & Estimated Damages** (positive correlation). This implies that more government funding to local governments can result in less damages by fire incidents while spending money on vehicles may not really help as much. We theorize that more vehicles correlate positively with more damages because vehicles are given mainly to metropolitan areas and the roads in these areas might not support the easy traversal of the fire vehicles. We recommend that further study could focus on the quality of the roads in these areas as to ascertain the reason that more vehicles do not necessarily help. Furthermore, the location of fire stations may also be a factor to consider since fires may sometimes occur in inaccessible areas. Furthermore, maintenance of fire trucks may also contribute to their capacity in lessening fire damages as the average acquisition age as seen in I is 16 years.

V. Conclusions and Limitations

Fires are a serious problem in the Philippines, and measures need to be done in order to minimize the amount of damage that an incident can cause. We can see that even though some areas are more urbanized than other locations, they actually tend to have more devastating fires than the more rural areas in the Philippines. This might be because the problem does not lie with a lack of resources on BFP's part, but other factors such as the previously mentioned population density and the amount of traffic in an area. Our tests show that additional funding from the government can help in mitigating the damages from fire incidents in an area. However, the IRA is not an accurate representation of the amount of funding that the fire department of various local governments receive, since the allocation of resources of these units might differ. The IRA is only meant to be a rough approximation. Additionally, the dataset we used only measures the scale of the damages of fire incidents based on the estimated damages in Philippine pesos, which might not be as accurate of a representation of the true damage that a fire causes. This is due to the disproportionate value of different property in the Philippines. For example, a fire that destroys an entire village in the without any injuries or deaths could have less estimated damages than when another different fire destroys a bank (also without any injuries or death). The value of property in different areas are also different. As stated in Rahim's paper [2], the methodology of the gathering of data for the dataset might also skew the data. Since the dataset only covers recorded fire incidents, care should be taken in considering if the more metropolitan areas just have a better record-keeping system than the non-metropolitan areas, which would result in more accurate data than those recorded in the provinces.

REFERENCES

[1] J. Mayuga, "Tragedy of fires: Death and destruction in the philippines," *Business Mirror*. [Online]. Available: https://businessmirror.com.ph/2018/03/21/tragedy-of-fires-death-and-destruction-in-the-philippines/

[2] M. S. N. A. Rahim, "The current trends and challenging situations of fire incident statistics," *Malaysian Journal of Forensic Sciences*, vol. 6, no. 1, pp. 63–78, 2015.

[3] S. Särdqvist and G. Holmstedt, "Correlation between firefighting operation and fire area: Analysis of statistics," *Fire Technology*, vol. 36, no. 2, pp. 109–130, May 2000. [Online]. Available: https://doi.org/10.1023/A:1015450308130

[4] J. Xin and C. F. Huang, "Fire risk assessment of residential buildings based on fire statistics from china," *Fire Technology*, vol. 50, no. 5, pp. 1147–1161, Sep 2014. [Online]. Available: https://doi.org/10.1007/s10694-013-0327-8

[5] A. Anderson and O. A. Ezekoye, "A comparative study assessing factors that influence home fire casualties and fatalities using state fire incident data," *Journal of Fire Protection Engineering*, vol. 23, no. 1, pp. 51–75, 2013. [Online]. Available: https://doi.org/10.1177/1042391512469521

[6] R. Nisanci, "Gis based fire analysis and production of fire-risk maps: The trabzon experienc," *Scientific Research and Essays*, may, volume= 2010.

[7] V. Pisarenko and M. Rodkin, *Heavy-Tailed Distributions in Disaster Analysis*, 01 2010, vol. 30.