

## THE DEVELOPMENT OF EARTHQUAKE VISUALIZATION DATA

### MAPPING IN INDONESIA PERIOD 2014 – 2016

HENDY TANDI & FRISKA NATALIA

*Universitas Multimedia Nusantara, Tangerang, Indonesia*

#### ABSTRACT

*The earthquake was one of the disasters that can give the effect of damage of the material or the soul of a very big if. Almost every two years once an earthquake-ravaged Indonesia is quite large and takes a lot of sacrifices, for example only the end of the year 2016, the province of Nangroe Aceh Darussalam again hit by quakes of 6.5 SR., consuming 102 fatalities and hundreds of people were injured. This is because 90% of the area in Indonesia has the potential to be affected by the earthquake happened. In this study an interactive dashboard was made with the aim to be able to provide information in the form of visualization of the earthquake that occurred in the territory of Indonesia, it is also helpful to show the magnitude of the strength, depth, intensity of earthquakes in the region of Indonesia. In this research, Visual Data Mining Method (VDM) is used in building a dashboard. The interactive dashboards created using Microsoft Power BI, Open Refine and Pentaho Data Integration which used for ETL processes (Extract, Transform, Load). The results obtained from the analysis of data visualizations that have been built shows the average area of Indonesia had the intensity of 3 to 5 MMI and the other results also showed an increase in the average strength of the earthquake at each year in Indonesia.*

**KEYWORDS:** *Data Visualization, Visual Data Mining (VDM), Dashboard, Earthquakes & BMKG.*

**Received:** Mar 20, 2018; **Accepted:** Apr 10, 2018; **Published:** May 08, 2018; **Paper Id.:** IJCEITRJUN20186

#### INTRODUCTION

Decision making in emergencies and preparation in the face of natural disasters is very important to be prepared. The use of information technology can be applied in different fields of research. In the study of earthquake level in all regions in Indonesia, then the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG) have been noted and provided a lot of data about the genesis of earthquakes that occur in the future are studied in order to Indonesia. However, the data that has been provided by BMKG is cannot be used directly by the people to get valuable information. In this study, the author intends to display the data mapping visualization results of the earthquake in Indonesia by utilizing data of earthquakes in Indonesia owned by BMKG.

The objective of the research is developing the interactive dashboard to show the areas in Indonesia that are most frequently to earthquakes or most affected are rarely affected by the earthquake. The visualization also will give an overview about the most frequent category of quake happens in Indonesia based on the Scale of Earthquake Intensity (SIG) scale as well as showing the high low strength, depth, and intensity of earthquakes that occur in the Indonesian. The dashboard can give a brief overview of the location of the occurrence of an earthquake, earthquake, earthquake occurrence time, the depth of the earthquake, complete with a comparison in the form of graphs. Previous related research is visual data mining method is a method that is suitable for visualizing large amounts of data and also not structured. In this study, Microsoft Power of BI is used to create a

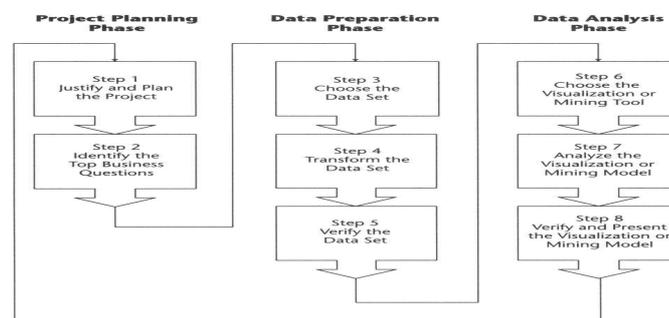
data visualization. Microsoft BI software is the Power with the ability to do data visualization in a comprehensive manner and change it in the form of graphs that are highly interactive and easy to understand (Rossi, 2005).

There are several references to research related to data visualization and Visual Data Mining methods. The first study is the research with aims to make it easier to understand and to see human data is not structured in large numbers. This study explains some of the goals or tasks in the use of the Visual Data Mining methods i.e. it allows users to explore the data in a large and looking for interesting information (Rossi, 2005). The second study has the purpose of classifying information used for data visualization based on the type of the data. This research revealed some data types that can be used for data visualization is a type of one-dimensional data such as data, two-dimensional data, multidimensional data, text, hypertext, hierarchies and graphs, and algorithms and software (Ward, 2002). The third study reveals a lot of information that can be conveyed to the listener or the users of visualization, for example by using line charts, tables, and graphics folders. By using the corresponding model visualization, information can be communicated by itself and provides benefits to its users (Ryan, 2015).

## PROBLEM STATEMENT AND METHODOLOGY

Visual Data Mining method is a method that provides the cycle to build a data visualization. The method which comprises several phases and in each phase, there are a few steps that makeup one unit of a cycle. This cycle can be started from the data preparation, data processing, data storage, and visualization.

There are three main phases in the method of visual data mining as described in Figure 1.



**Figure 1: Visual Data Mining Phase and Step**

In Project Planning Phase there are two main steps:

- Plan and Justify the Project, in this step, there are four actions which must be identified: planning the visualizing, how to make the visualization, what's the purpose of the visualization and determine the resources are needed in this research. This research requires resources such as earthquake data and tools that will be used in visualizing data, such as Microsoft Power of BI.
- Identify the Top Business, in this step, is intended to determine the needs of the main users of visualization. Based on the interview, the result is acquiring data about the region wherein Indonesia frequently suffered damage of earthquakes, in addition, the author also obtains technical information about earthquakes and earthquake category Division also based on intensity.

In Data Preparation Phase, there are three steps to prepare the necessary data, in addition to the use of tools and procedures for data processing also occurs in this phase.

- Choose the DataSet, data collection was taken through the official website online data belonging to BMKG Indonesia.
- Transform the Data, the next set of data in a format that Microsoft Excel will process the cleaned data. The cleaning and Extract, Transform, Load (ETL) process is using Open Refine. Open the Refine is one of data cleaning software created by Google and run on the Windows operating system. Google is designing software for use in accessing data from a number of small to large, clean and change data, and perform a reconciliation and comparison data.
- Verify the DataSet, after the data has been cleansed and transformed, then the checking process to make sure the data is completely free from bias and error. Pentaho is used in this verification process. Pentaho can do a lot of the process of data integration as a whole including ETL process and verification of data. The advantages of Pentaho compared to other similar software is the dashboard, which displayed by the Pentaho Graphic User Interface and looks more user-friendly, in addition, Pentaho can also be integrated with other business intelligence platform. Therefore, it is easier to use on the scope of cross-platform.

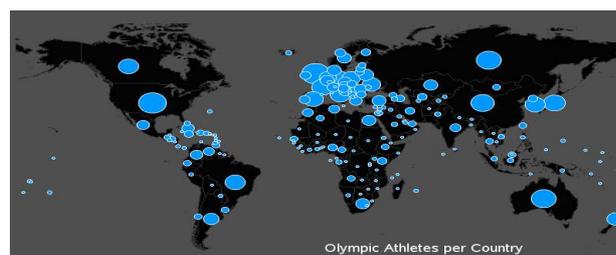
In Data Analysis Phase, there are three steps to visualize the data. In this phase the data also started visualized, analyzed and presented the results.

- Choose the Visualization or Mining Tools, in this step the comparison of tools is done. Based on that, the tool that used in this research is Power Business Intelligence (BI). Power BI is a product of Microsoft technology company in which Microsoft products such as Microsoft Office, Microsoft excel, and access can be accessed easily by Microsoft Power BI, in addition, there are also a lot of visual customization that brought to Microsoft Power BI.
- Analyze the Visualization or Data Mining Models, in this step the analysis of the data visualization has been formed. From the selected data, the shape data visualization can be analyzed to make sure the information displayed on each of the graphs that have been formed.
- Verify and Present Analysis, at the last stage of this step the conclusions of the information have been obtained from the previous step (analysis phase). This information is subsequently verified back righteousness and also presented to major users who need it, in this case, it is BMKG Indonesia.

## **ANALYSIS AND RESULTS**

### **Project Planning Phase**

In this research, the visualization of data that will be made in the form of graphs and the folder is created in the form of visualization of the map which described in Figure 2.



**Figure 2: Map Graphics**

Folders allow users easier to understand the meaning of the data in certain regions. The earthquakes data from 2014-2016 was used in this research by recommendation of the Agency for meteorology, climatology, and Geophysics Indonesia, due to comparing mapping earthquake in Indonesia needed earthquake data more than a year. Therefore, BMKG suggested taking earthquakes data over the period of the last 3 years the year 2014, 2015, and 2016. In this study will use that data which provided directly by the Agency for meteorology, climatology, and Geophysics Indonesia as follows:

- The Data of earthquakes throughout the region Indonesia 2014 year period up to 2016.
- Earthquake Data that can be felt over the years 2014 up to 2016.
- Earthquake Data the province of Nangroe Aceh Darussalam 2014 year period up to 2016.

In this research there is some important information which BMKG needed in the interactive dashboard:

- Display the earthquake data on a specific area or a specific color information that shows an information for its users.
- Describe the average maximum strength, depth, and intensity of each month in a year.
- Can view mapping distribution of earthquakes that have occurred in certain areas. This visualization needed for easy viewing of the information contained in the region.

### Data Preparation Phase

Based on the results of the interviews have been conducted and the results of the request data provided by BMKG, then the data that will be used are as follows:

Data of earthquakes throughout the region in Indonesia (in 2014, 2015, 2016), with details as follows :

**Table 1: Database of Earthquake Data for all Regions in Indonesia**

Column	Type	Description
Date	General	The date of the occurrence of the earthquake
Time	General	The start time of occurrence of earthquake
Latitude	General	Point latitude coordinate earthquake
Longitude	General	Coordinates longitude point earthquake
Depth	General	The depth of the earthquake (KM)
Mag	General	The strength of an earthquake
TypeMag	General	The type of Earthquake happens
smaj	General	The distance of the stations farthest away from epicenter
smin	General	The distance of the nearest station from the epicenter
az	General	The greatest distance between stations on the note taker
rms	General	Error rate calculation
cPhase	General	The number of the station's earthquake
Region	General	The location of the earthquake

Data of the earthquake was felt (in 2014, 2015, 2016), with details as follows :

**Table 2: Database of the Data Earthquake was Felt**

Column	Type	Description
Date	General	The date of the occurrence of the earthquake
Time	General	The start time of occurrence of earthquake
Latitude	General	Point latitude coordinate earthquake
Longitude	General	Coordinates longitude point earthquake
Depth	General	The depth of the earthquake (KM)
Mag	General	The strength of an earthquake
Region (MMI)	General	The location of the earthquake A scale to measure earthquake intensity

Data of the earthquake in the province of Nangroe Aceh Darussalam

**Table 3: Database of Earthquake Data Columns of the Province of Nangroe Aceh Darussalam**

Column	Type	Description
Date	General	The date of the occurrence of the earthquake
Time	General	The start time of occurrence of earthquake
Latitude	General	Point latitude coordinate earthquake
Longitude	General	Coordinates longitude point earthquake
Depth	General	The depth of the earthquake (KM)
Mag	General	The strength of an earthquake

Cleaning data is needed to simplify and shorten the time of transformation phase of workmanship by using Open Refine. After that, from the data has been cleaned need to change into Mysql. In order to operate this tools, need to understand the GREL (General Refine Expression Language), which is the language that is used as the syntax for modifying data on gross Open Refine. The first phase of the cleanup data does is remove the string behind the information. Examples include word WIT, Km, SR., and MMI. This step is very important because the data types such as integer and string can't be lumped together in a single cell in Mysql. Figure 3 describes an example of the data that has been cleaned.

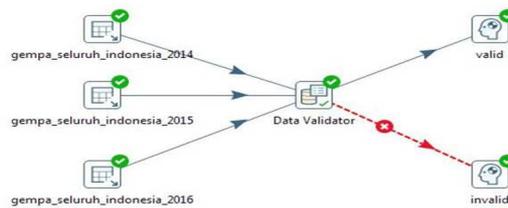
250 rows		Extensions			
Show as: rows records		Show: 5 10 25 50 rows per page 1 - 10 next last			
All	NO	TANGGAL	WAKTU GEMPA	LINTANG	
1.	1	2014-01-03T00:00:00Z	20:06:34 WIB	0.15 LS	
2.	2	2014-01-05T00:00:00Z	10:42:12 WIB	0.69 LS	
3.	3	2014-01-09T00:00:00Z	19:05:48 WIB	0.88 LS	
4.	4	2014-01-09T00:00:00Z	23:16:48 WIB	7.12 LS	
5.	5	2014-01-18T00:00:00Z	07:19:55 WIB	2.5 LS	
6.	6	2014-01-18T00:00:00Z	09:13:31 WIB	1.97 LS	
7.	7	2014-01-18T00:00:00Z	11:05:38 WIB	2.13 LS	
8.	8	2014-01-20T00:00:00Z	11:45:14 WIB	4.94 LS	
9.	9	2014-01-20T00:00:00Z	23:27:29 WIB	2.41 LS	
10.	10	2014-01-22T00:00:00Z	13:00:17 WIB	2.07 LS	

**Figure 3: Data Cleaning Process**

**Table 4: The End Result of Earthquake Data Throughout the Region in Indonesia**

Column	Type	Example of Data
Date	General	01/01/2014
Time	General	03:20:13.7
Latitude	General	-6.99
Longitude	General	127.71
Depth	Number	252
Mag	Number	4,2
Region	General	Banda Sea

All data on the research that has been cleaned on the previous stage will be verified again using a validator function at Pentaho which showed in Figure 4.



**Figure 4: Data Verification Process**

The validation is running and then Pentaho will give final conclusions upon the validation is done, where will the appearance of the data is valid or invalid in Figure 5.

**Execution Results**

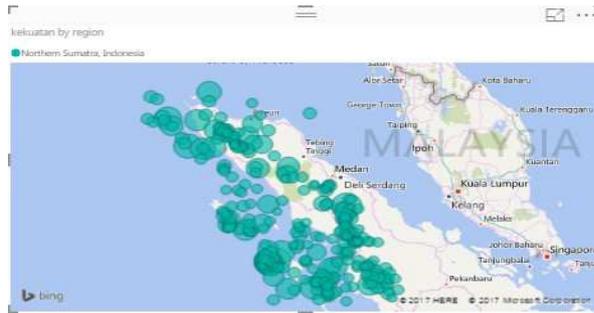
Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active
1	gempa_seluruh_indonesia_2014	0	0	6668	6668	0	0	0	0	Finished
2	gempa_seluruh_indonesia_2016	0	0	4782	4782	0	0	0	0	Finished
3	gempa_seluruh_indonesia_2015	0	0	5073	5073	0	0	0	0	Finished
4	Data Validator	0	16523	16523	0	0	0	0	0	Finished
5	valid	0	16523	16523	0	0	0	0	0	Finished
6	invalid	0	0	0	0	0	0	0	0	Finished

**Figure 5: Execution Process**

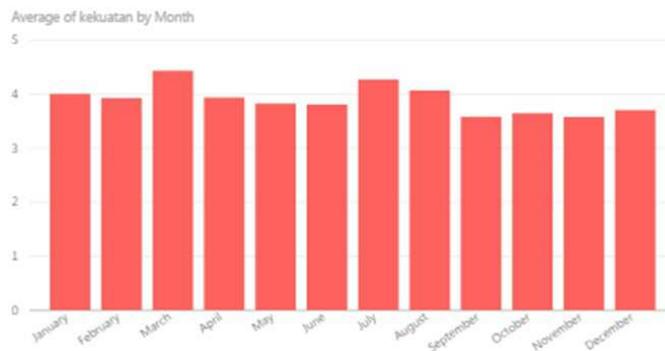
**Data Analysis Phases**

Visualization tools used in this research is Microsoft Power BI. The tool can provide advantages rather than some other visualization tools. The advantages i.e. like visualization and dashboards that are more interactive, the availability of visualization in the form of mobile, simply connect to Power BI with other Microsoft products. Visualization in this study consists of a map graphic, a clustered column chart, Line chart, and data tables. The map graphic in Figure 6 is being used to show the mapping of an earthquake on the map areas in Indonesia.



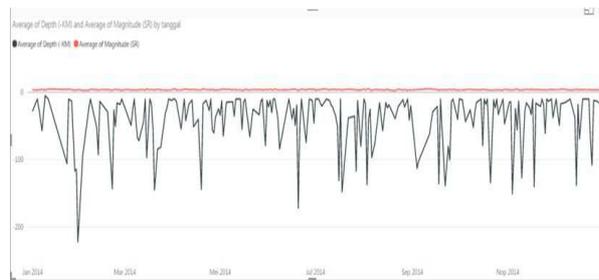
**Figure 6: Map Graphic Model**

A clustered column chart in Figure 7 is used to show a high-value low strength, depth, and intensity of the earthquake that occurred in each month.



**Figure 7: Clustered Column Chart Model**

A Line chart used in this research to show how the earthquake with the strength, depth, and intensity. Line Charts are also used to show the strength, depth, and intensity every day which described in Figure 8.



**Figure 8: Line Chart Model**

The data table in Figure 9 is used to show the earthquake data when selected on a folder, such as the strength, depth, intensity, and the name of the region.



**Figure 9: Data Tabel Model**

The next step is to analyze the visualization that has been built in Figure 10 until Figure 13. The analysis is done by using dashboard visualization that has been made: Dashboard visualization of earthquake data throughout Indonesia from 2014 to 2016, Earthquake visualization dashboard in Province of Nanggroe Aceh Darussalam from 2014 until 2016, and Dashboard visualization of data of earthquake felt throughout Indonesia from 2014 to 2016, as well as the earthquake visualization dashboard perceived by Nanggroe Aceh Darussalam Province from 2014 to 2016.

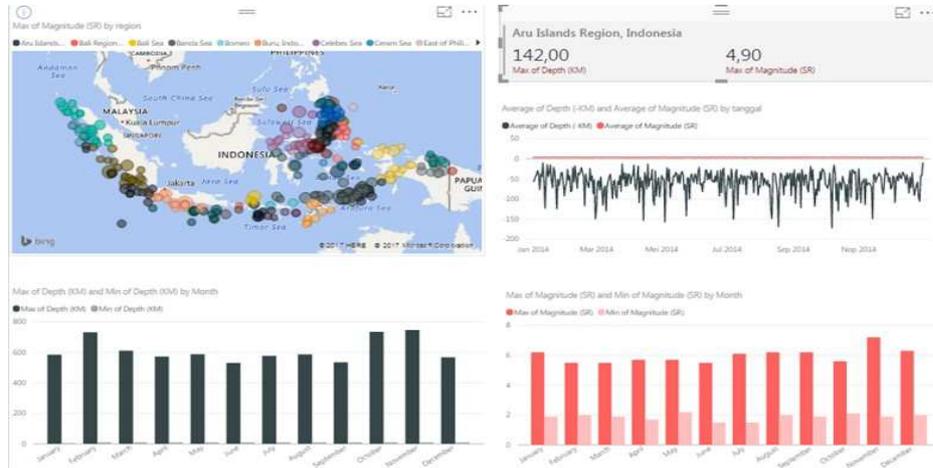


Figure 10: The Earthquake Data Visualization Dashboard across Indonesia

The first analysis performed on dashboard visualization of earthquake data throughout Indonesia from 2014 to 2016 (Figure 10), it can be concluded that the average maximum strength of earthquake in Indonesia is 6,123 SR with the average maximum depth of 605,4 KM per year. From the average maximum strength of earthquakes throughout Indonesia that occurred in the last 3 years showed an increase: in 2014 of 5.96 SR, in 2015 of 6.16 SR, and in 2016 at 6.25 SR.

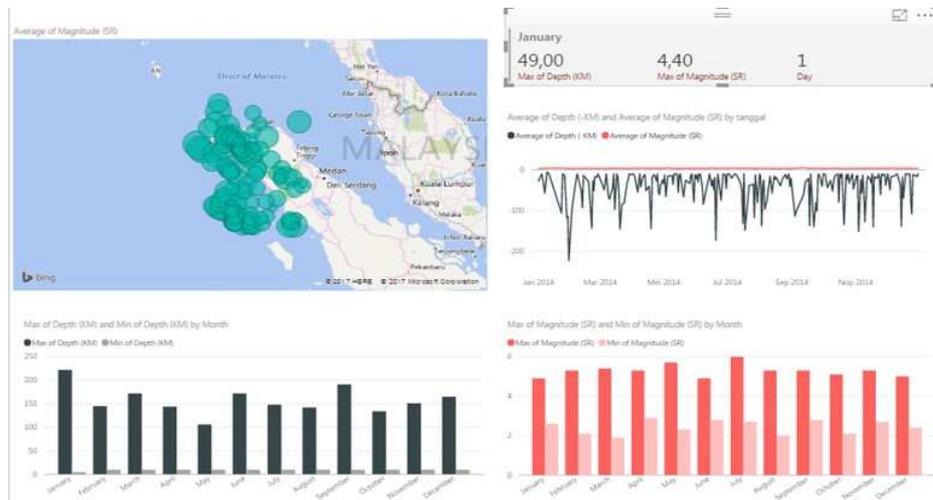
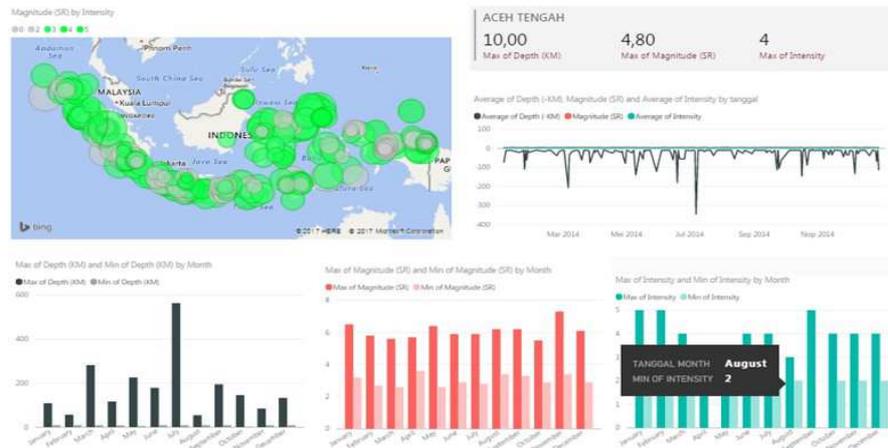


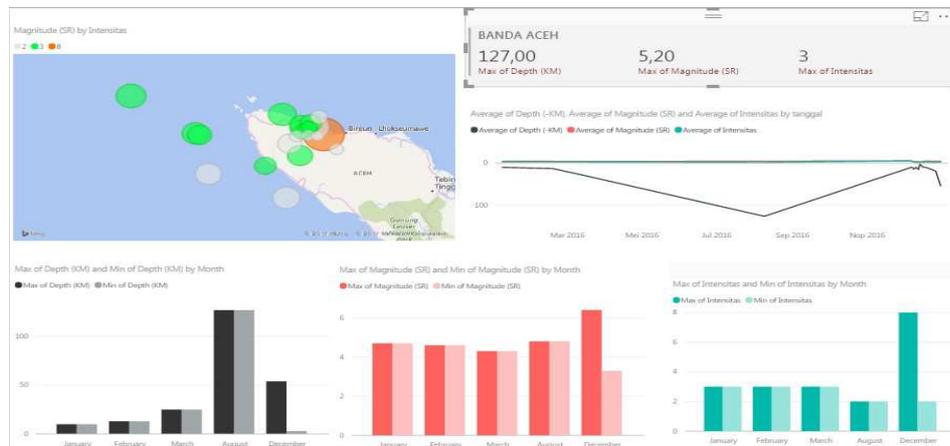
Figure 11: The Earthquake Visualization Dashboard in Province of Nanggroe Aceh Darussalam from 2014 Until 2016

The second analysis was conducted on earthquake visualization dashboard in Nanggroe Aceh Darussalam Province from 2014 until 2016 (Figure 11), which concluded that the maximum mean of earthquake strength in Nanggroe Aceh Darussalam Province is 5,347 SR with an average of 142,1 KM depth. At Province Nanggroe Aceh Darussalam also experienced an increase in earthquake strength: 2014 amounted to 5.29, then increased in 2015 by 5.36 SR, and recovered in 2016 amounted to 5, 39 SR.



**Figure 12: The Earthquake Data Visualization Dashboard of Earthquake Felt Throughout Indonesia from 2014 To 2016**

The third analysis conducted on the dashboard of earthquake data visualization is felt throughout Indonesia 2014 to 2016 gives the result described in Figure 12. The average maximum strength of the earthquake in Indonesia is 6.26 SR with an average depth of 208.5 KM and average maximum intensity of 4.403, which entered into the perceived category. In this visualization also experienced an increase in the average maximum strength of the earthquake: in 2014 at 6.08, then increased by 2015 by 6.31, and increased again in 2016 at 6.39.



**Figure 13: The Earthquake Data Visualization Dashboard Perceived by Nanggroe Aceh Darussalam Province From 2014 To 2016**

The fourth analysis conducted on the earthquake visualization dashboard was felt by the Province of Nanggroe Aceh Darussalam from 2014 to 2016 (Figure 13). In this analysis, the maximum average of earthquake strength in Nanggroe Aceh Darussalam Province also increased, wherein 2014 recorded the maximum average strength the earthquake in Nanggroe Aceh Darussalam Province amounted to 5.12 SR, then increased by 2015 by 5.29 SR, and increased again in 2016 of 5.65 SR. Also, note that Nanggroe Aceh Darussalam Province is the only province in Indonesia that has an earthquake intensity of 8 MMI over the past 3 years with moderate damage status.

From the fourth analysis, concludes that there is an increase in earthquake strength over the last 3 years (2014 until 2016) with an average increase in strength of 0.31 SR per year. The following is more detailed information:

**Table 5: Calculation of Earthquake Data Visualization Dashboards Across Indonesia Year 2014 Year Up To 2016**

Desc	2014				2015				2016				Average			
	Max	Min	H	L	Max	Min	H	L	Max	Min	H	L	Max	Min	H	L
Depth	614,08	9,58	747	5	579,58	9,33	92	2	622,41	9,25	750	1	605,4	9,39	529,67	2,67
Magnitude	5,96	1,87	7,19	1,5	6,16	2,08	6,9	1,39	6,25	1,76	7,8	1,5	6,123	1,9	7,2967	1,46

To conduct a more in-depth analysis, one of the cities in Nanggroe Aceh Darussalam Province, Banda Aceh, is chosen as the city to be analyzed next. Banda Aceh city was chosen because it is one of the most experienced cities in Nanggroe Aceh Darussalam Province in a year. In 2014 there are 12 earthquakes in the city of Banda Aceh with an average earthquake measuring 4.8 SR, then in 2015 the city of Banda Aceh back 12 times the earthquake with an average earthquake 5.0 SR, and last in 2016 city Banda Aceh experienced 6 quakes with an average earthquake strength of 4.8 SR. This proves the average strength of the earthquake that occurred in the region of Indonesia ranges from  $\pm 5$  SR.

The average result of the maximum strength and depth between analysis 1 and 2 with 3 and 4 is different due to the difference of data source, where the data in analysis 1 and 2 are taken from BMKG Indonesia which records all earthquakes in Indonesia without any minimum strength or in analysis data 3 and 4 were taken from Ina TEWS (Indonesia Tsunami Early Warning System) which only recorded earthquakes with a minimum strength of 2.09 SR. From these two data sources give the calculation of the maximum average strength of earthquake in Indonesia and Nanggroe Aceh Darussalam Province, that is with the average strength of  $\pm 5$  to 6 SR earthquake.

### Analysis Results

The analysis was conducted to see the effect of the condition before and after the visualization of this earthquake mapping dashboard showed in Table 6. Below is a table of analysis results on the comparison of conditions before and after the visualization of earthquake mapping dashboard visualization.

**Table 6: Comparison Analysis**

Aspects	Before	After
Imaging or plotting the point of the earthquake that occurred in the territory of Indonesia	Cannot see the point of the earthquake that occurred in the territory of Indonesia	Imaging or plotting of earthquake points by region can be seen with the map graphic.
Comparison of strength, depth, and intensity of earthquakes	The strength, depth, and intensity of earthquakes can not be known magnitudes	The strength, depth, and intensity of earthquakes now can be compared
Earthquake point imaging is based on the color of the BMKG SIG scale	Can not see earthquake color from GIS scale in the form of earthquake point	The color of earthquakes of GIS scale can be seen at the point of earthquake using map graphic
Tool in socialization and education to the community about the earthquake by BMKG	Lack of tools to educate people about earthquakes	BMKG can utilize this visualization as a tool to educate the public

The benefits of this visualization are imaging of earthquake points in map graphics, strength, depth, and intensity of earthquakes in the Clustered Column Chart Model, and also provide color-based imagery at the point of earthquakes using map graphics.

## CONCLUSIONS

From the analysis it can be concluded that the visualization that has been made has given some benefits such as can already see the imaging or plotting point of earthquake based on its region by using map graphic contained in the dashboard, comparison of information such as strength, depth, and intensity of earthquake also have viewed on the dashboard, in addition to the map graphic in the dashboard also can see the specific colors based on SIG Scale. This visualization can also be a tool that can be used to educate and socialize the community about earthquakes by BMKG as its users. The visualizations built in this research have been published into the Microsoft Power BI website.

## REFERENCES

1. Card, S. (1998). *Readings in Information Visualization : Using Vision to Think*. Morgan Kaufmann.
2. Friedman, V. (2008). *Data Visualization and Infographics*. Monday Inspiration.
3. Keim, D. A. (2002). *Information Visualization and Visual Data Mining*. 5-8.
4. Rossi, F. (2005). *Visual Data Mining and Machine Learning*. 5.
5. Soukup, T. (2002). *Visual Data Mining-Techniques and Tools for Data Visualization and Mining*. New York: Wiley.
6. Ward, M. (2002). *Visual Data Mining Techniques*. 3-11.
7. Connolly. (2005). *A Practical Approach to Design, Implementation, and Management*. Amerika: Longmen.
8. Rahm, E. (2000). *Data Cleaning: Problems and Current Approaches*. Jerman.
9. Scheps, S. (2008). *Business Intelligence For Dummies*. US: Willey.
10. Turban, E. (2007). *Decision Support and Business Intelligence System*. New Jersey: Pearson Prentice Hall.

