

PAPER • OPEN ACCESS

Earthquake data analysis and data visualization of Maharashtra state, India from 1912 to 2009 using R programming

To cite this article: Vaishnavi Battul *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **989** 012029

View the [article online](#) for updates and enhancements.



ECS **240th ECS Meeting**
Oct 10-14, 2021, Orlando, Florida

**Register early and save
up to 20% on registration costs**

Early registration deadline Sep 13

REGISTER NOW

Earthquake data analysis and data visualization of Maharashtra state, India from 1912 to 2009 using R programming

Vaishnavi Battul^{1,3} Helen Santhi M¹ and Malathi G²

¹School of Civil Engineering, Vellore Institute of Technology, Chennai, TamilNadu, India

²School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, Tamil Nadu, India

³Email: vaishnavi.2019@vitstudent.ac.in

Abstract. This paper presents the seismic data analysis and visualization using R programming for the state of Maharashtra, India. Maharashtra has witnessed a significant number of earthquakes ranging from low to high in magnitude. The past earthquake data of a particular location are useful to the civil and structural engineers for the safe and efficient planning, designing and construction of seismic resistant buildings in that locality. In this paper the past earthquake data of Maharashtra from the year 1912 to 2009 are collected and analyzed. There are 53 earthquake data with magnitudes ranging from 3.6 to 7.2. This analysis shows the relationship between magnitude and depth, intensity and PGA, visualization of classification of earthquakes depends on magnitude and intensity.

1. Introduction

Data analysis and data visualization are being used in almost all fields of science, engineering and business. In order to understand and make conclusions on any project, interpretation of data is the key function which majorly depends on how efficiently the data analysis and visualization are carried out. Many researchers have started analyzing earthquake data worldwide [1, 4, 5] and contributing towards the design of seismic resistant infrastructure. Earthquake is one of the major disasters threatening the world and the assessment of risk involved is to be evaluated to protect the life of people and their possessions. In India, most of the parts are vulnerable to earthquake and the loss of life and damages to the infrastructure are huge in some of the past earthquakes. The ill effects of earthquakes can be minimized by understanding the characteristics of earthquakes using the past seismic history of any particular region in the country. The seismic data available can be analysed [7, 9, 10] visualized [2, 3, 6, 8] and interpreted to protect the life and to reduce the damages to the infrastructure from future earthquakes using sophisticated computer tools. In this paper, data analysis and data visualization of earthquake data between 1912 and 2009 in Maharashtra State, India is carried out using R software. The graphical representation of the earthquake data in the study area gives better understanding of the site specific characteristics of the past earthquakes.

2. Data analysis and data visualization with ggplot using R

Following are the packages and libraries that are loaded to attain useful outputs related to seismic activity in Maharashtra.

```
install.packages("dplyr")
install.packages("ggmap")
install.packages("ggplot2")
library(dplyr) library(ggmap)
library(ggplot2)
```



2.1 Maharashtra earthquake dataset

The past earthquake data of Maharashtra State are collected from the year 1912 to 2009 and the following steps are followed to read and view the data (Figure 1).

Input:

```
maha <- read.csv("maha.csv", stringsAsFactors = FALSE)
```

summary() function

Input: summary(maha)

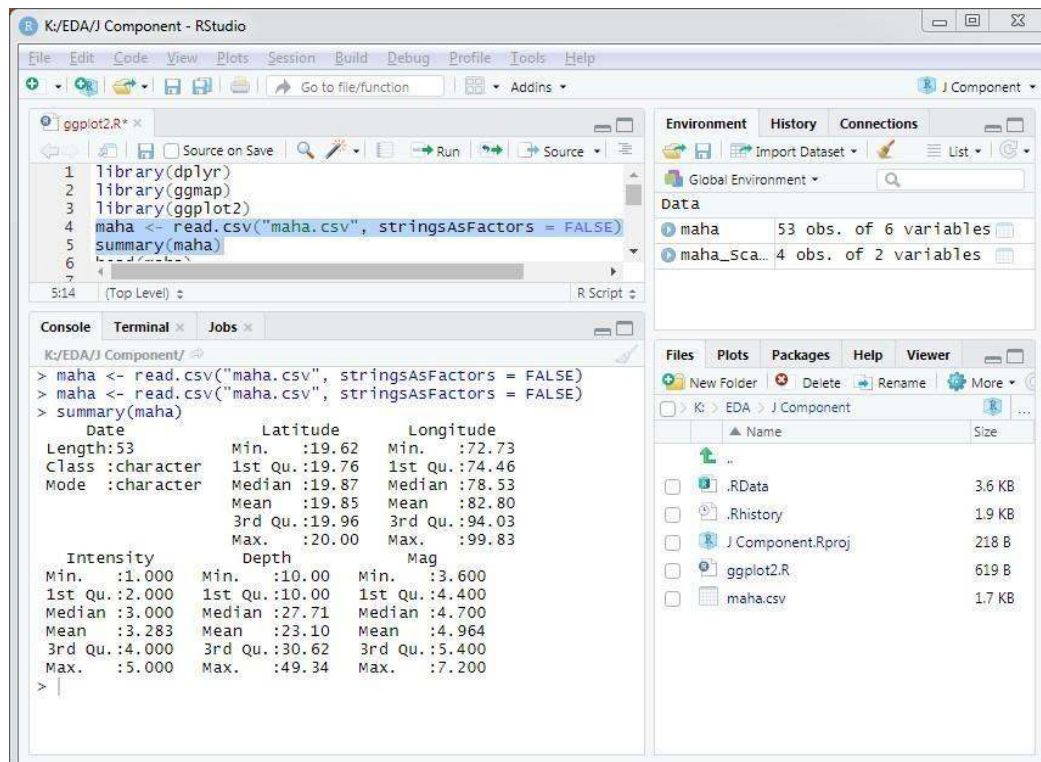


Figure 1. Data downloaded ready to read and view.

head() function

- Input:
head(maha)

It displays the data in Tabular form with headings to the columns

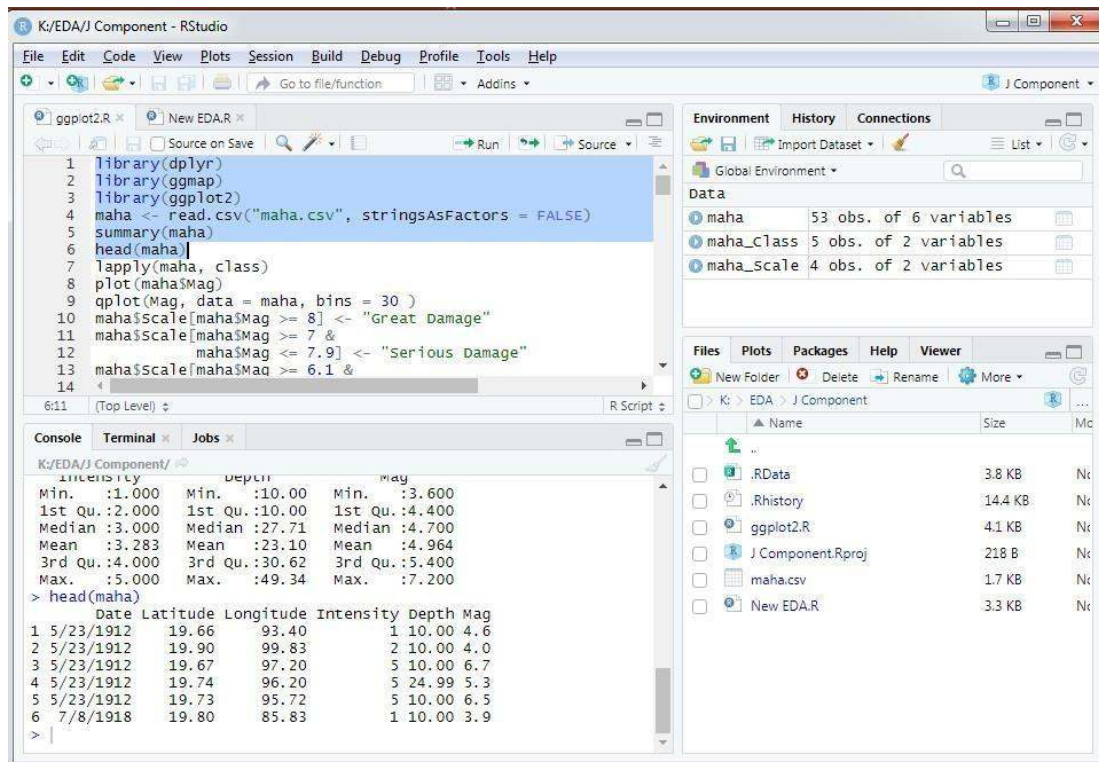


Figure 2. Data arranged in columns.

`lapply()` function

- To display the earthquake data the above outputs from the `summary()` and `head()` functions are helpful.
- The `lapply()` command is useful to see classes of data.
- Input
`lapply(maha, class)`

2.2 Maharashtra earthquake plot

Figure 3 depicts the plot of the magnitude of each earthquake. It is seen that the plot is scattered because of the frequency of occurrence of earthquakes. From the plot it is recognized that most of the earthquakes are lower scaled ones.

Input

```

#Plot of Magnitude of each
Earthquake plot(maha$Mag)

```

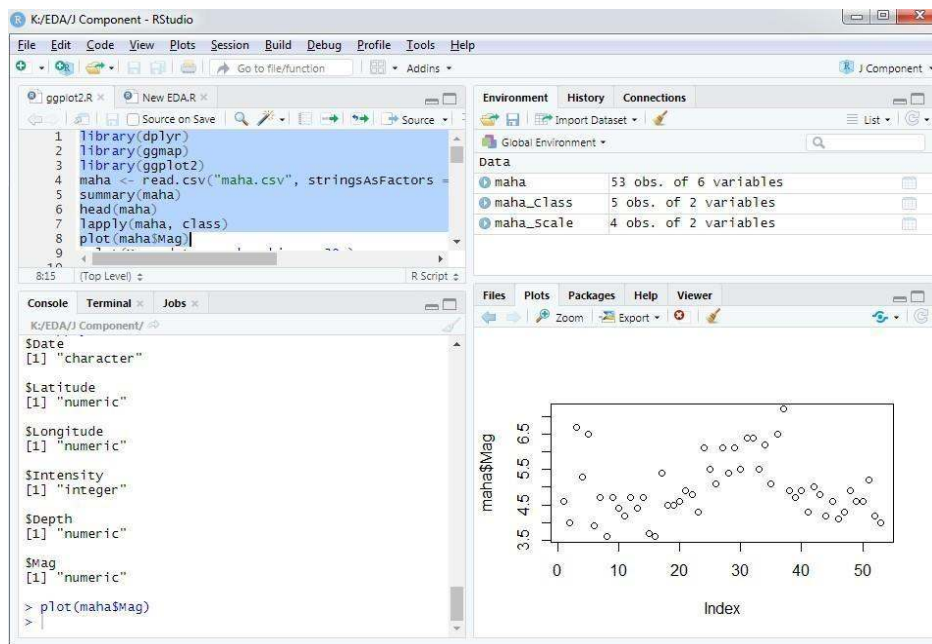


Figure 3. Earthquake magnitude plot (Scattered).

The above details can also be presented in another way of plotting as shown Figure 4 that would be easily understandable, `qplot()`.

#qplot of Magnitude and Total Count
`qplot(Mag, data = maha, bins = 50)`

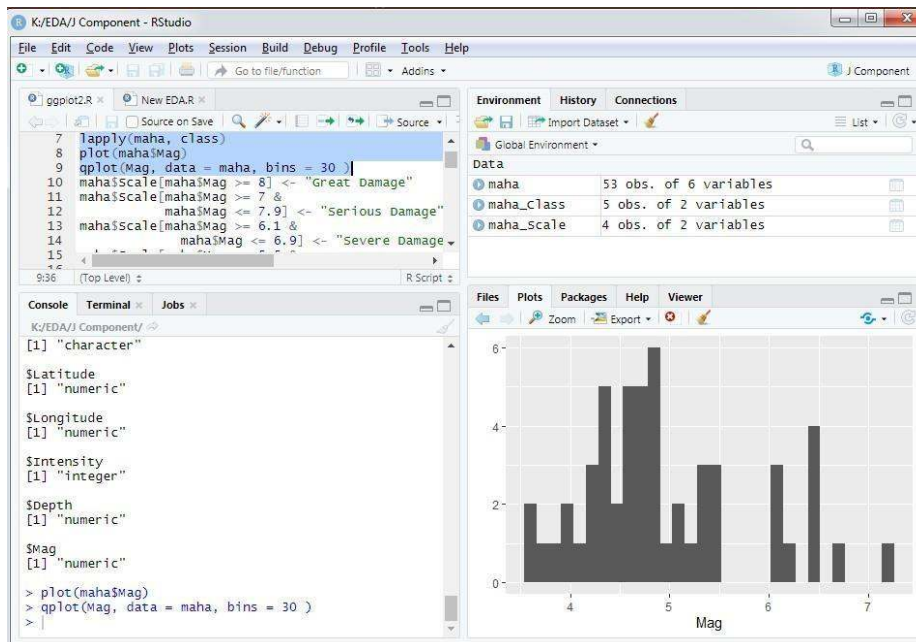


Figure 4. Earthquake magnitude plot (Bar chart).

2.3 Scale of earthquake magnitude

Generally the earthquake magnitude is scaled based on the severity of the earthquakes. Table 1 shows the various levels of earthquake scale adopted in the study and it is found that the damage level falls under minor damage.

Table 1. Damage level for different earthquake magnitude.

Magnitude	Damage level
< 2.5	Limited damage represents very small damage
2.5 to 5.4	Minor damage represents small damage
5.5 to 6.0	Slight damage represents noticeable damage
6.1 to 6.9	Severe damage represents highly noticeable damage
7.0 to 7.9	Serious damage represents very severe damage
>8.0	Great damage represents very high damage

The damage levels of earthquake in Maharashtra state can be visualized using the following steps and the plot is shown in figure 5.

```
ggplot(data = maha_Scale, aes(x = Scale, y = n, fill = Scale)) +
geom_bar(stat = "identity") +
xlab("Earthquake Magnitude by Scale")
+ ylab("Total Earthquakes") +

ggtitle("Maharashtra Earthquake Magnitude Scale") +
theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
theme(plot.title = element_text(hjust = 0.5))
```

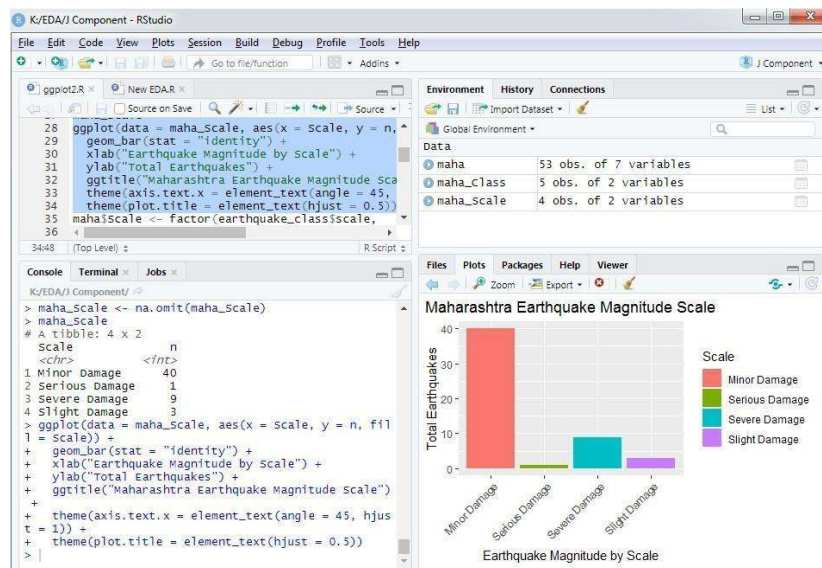


Figure 5. Visualization of earthquake damage levels.

2.4 Earthquake Scatter plot of magnitude and depth

Earthquakes represented in terms of magnitude and depth would be useful to check the influence of depth on magnitude of any earthquake. Here the earthquakes are classified based on the depth of origin of earthquake as follows. If the depth is from 0 to 70 km, it is shallow earthquake; from 70 to 300 km, it is intermediate earthquake; from 300 to 700 km, it is deep earthquake.

In order to find the influence of depth on the earthquake magnitude, a scatter plot is produced. Here the magnitude is scaled between zero and eight. From the plot, it can be seen that the correlation between the depth and the magnitude is positive.

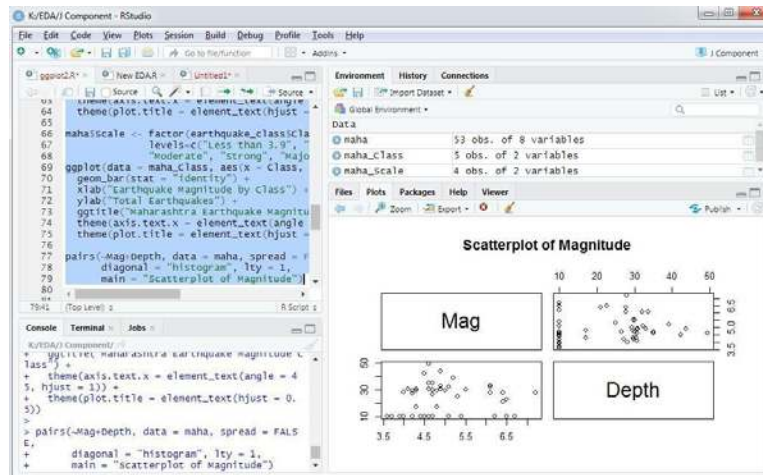


Figure 6. Depth versus magnitude plot.

2.5 `get_map()` and `ggmap()` for mapping earthquakes

Now the data is ready and for mapping, `get_map()` and `ggmap()` functions are used.

```
(map <- get_map(c(left = 72.25345, bottom = 15.321458, right = 100.099334, top = 21.559652)))
ggmap(map)
```

```
ggmap(map, extent = "device")
+ geom_point(data = maha,
aes(x = Longitude, y = Latitude, color = Mag, size = Depth)) +
scale_colour_gradient(low = "yellow", high = "red")
```

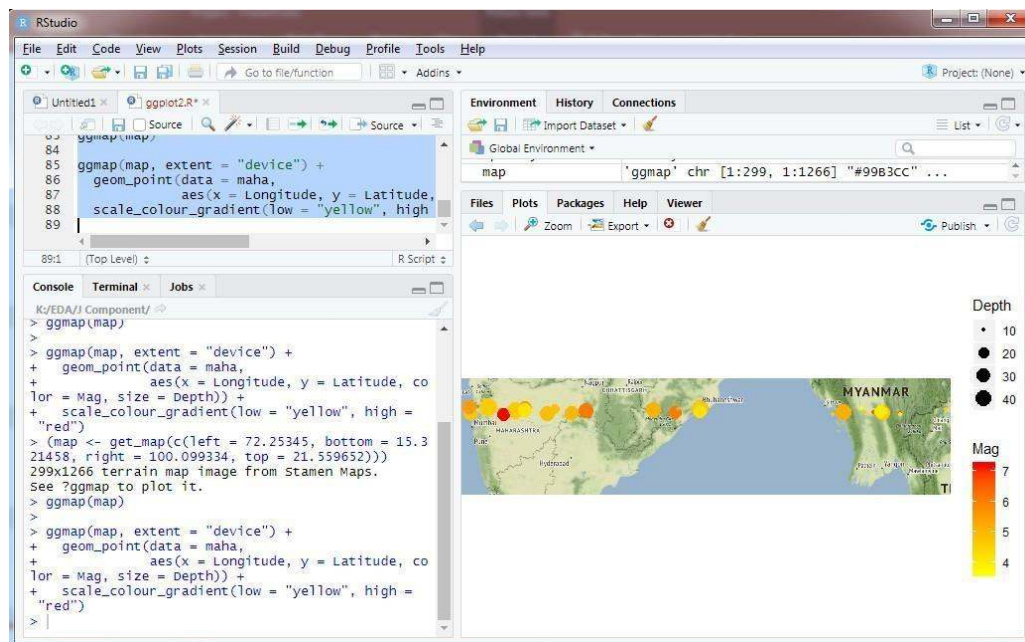


Figure 7. Mapping of earthquakes.

3. Regression model

Multiple regression analysis is carried out with three independent variables such as epicentral distance, depth and magnitude of earthquake and peak ground acceleration (PGA) as dependant variable. From the intercept and coefficient values, the mathematical equation is obtained as below.

$$Y = 1.963736194 \times 10^{-3} X_1 - 2.874770724 \times 10^{-4} X_2 + 7.174167726 \times 10^{-1} X_3 - 1.639089517 \times 10^{-1}$$

$$R^2 = 0.8054246502$$

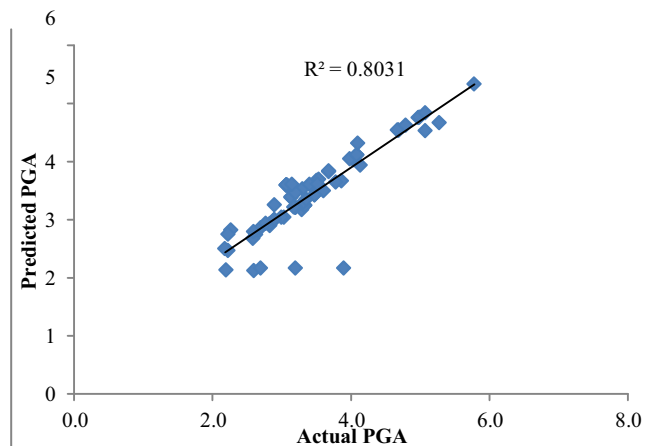


Figure 8. Correlation between actual and predicted PGA

4. Conclusion

The past earthquake data of Maharashtra state are collected, analyzed and visualized using R programming to understand the earthquake activities easily and meaningfully. Also the equation for PGA is arrived at from multiple regression analysis to predict the earthquake size. Similar kind of investigation can be conducted for all the vulnerable regions in the country to predict the characteristics of earthquakes and to design safer infrastructure.

5. References

- [1] Admasu F and Toennies K 2010 Anisotropic 3D seismic features for robust horizons correlation across faults *Proceed.IEEE Inte. Con. Image Pro.* (Carolina USA) 221
- [2] Chourasia A, Cutchin S and Aagaard B 2008 Visualizing the ground motions of the 1906 San Francisco earthquake comp *Geosciences* **34(12)** 1798
- [3] Li W and Wang D 2012 Visualization system of massive 2D seismic data *Inform.Engg.Appl* **15(4)** 748
- [4] Liu R C, Guo H Q and Yuan X R 2014 Seismic structure extraction based on multi-scale sensitivity Analysis *J. Visual.***17 (3)**157
- [5] Ma R A, Zhang A H and Yang J Y 2005 A fast volume rendering method for 3D seismic volume data *J.Com.-Aided Des.& Com. Gra.***17(5)** 1008
- [6] Naka T, Yamada M, Endo M, Miyazaki S and Hasegawa J 2006 Visualization of seismic-center distribution data for earthquake prediction Proc. of Nicograph Intl
- [7] Patel D, Giertsen C, Thurmond J, Gjelberg J and Groller M E 2008 The seismic analyzer-interpreting and illustrating 2D seismic data *IEEE Trans. Visual.Com.Gra.***14 (6)** 1571
- [8] Tung-Ju Hsieh, Cheng-Kai Chen and Kwan-Liu Ma 2010 Visualizing field-measured seismic data Proc. of IEEE Pacific Visualization Symposium 65
- [9] Yin L, Li XL, Zheng W, Yin Z, Song L, Ge L and Zeng Q 2019 Fractal dimension analysis for seismicity spatial and temporal distribution in the circum-pacific seismic belt *J.Earth Sys. Sci.***128 (1)** 22
- [10] Yu Y, Jia Z, Cao L, Zhao J D, Liu Z W and Liu J L 2018 Fast density-based clustering algorithm for location big data *J. Soft.* **29 (8)** 2470