

Project 2 – Major Earthquakes

MTH 161 – Fall 2024

Madison Griffith

2024-11-10

The data set that I have chosen to use is provided by OpenIntro.

Part 1: choose a dataset and propose a question

```
read_csv("https://www.openintro.org/data/csv/major_earthquakes.csv")
```

Rows: 123 Columns: 7

-- Column specification -----

Delimiter: ","

chr (3): month, area, region

dbl (4): year, day, intensity_richter, deaths

i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show_col_types = FALSE` to quiet this message.

A tibble: 123 x 7

	year	month	day	intensity_richter	area	region	deaths
	<dbl>	<chr>	<dbl>	<dbl>	<chr>	<chr>	<dbl>
1	1902	April	19	7.5	Quezaltenango and San M~	Guate~	2000
2	1902	December	16	6.4	Uzbekistan	Russia	4700
3	1903	April	28	7	Malazgirt	Turkey	3500
4	1903	May	28	5.8	Gole	Turkey	1000
5	1905	April	4	7.5	Kangra	India	19000
6	1906	January	31	8.8	Esmeraldas (off coast)	Ecuad~	1000
7	1906	March	16	6.8	Chia-i	Taiwan	1250
8	1906	April	18	7.7	San Francisco	Unite~	3000
9	1906	August	17	8.6	Valparaiso	Chile	3882

```
10 1907 October      21                8.1 Central                Asia      12000
# i 113 more rows
```

```
csv_table <- read_csv("https://www.openintro.org/data/csv/major_earthquakes.csv")
```

```
Rows: 123 Columns: 7
```

```
-- Column specification -----
```

```
Delimiter: ","
```

```
chr (3): month, area, region
```

```
dbl (4): year, day, intensity_richter, deaths
```

```
i Use `spec()` to retrieve the full column specification for this data.
```

```
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

The data set I have chosen to examine is Major Earthquakes. My research question is does the month affect the magnitude of the intensity richter?

::: Review the Data

Let's analyze a boxplot of the months, and see where the trend line is for the intensity richter. An intensity richter is a 10 logarithmic scale that analyzes the intensity of each earthquake. We will separate the data in each month and see if there is a difference in the intensity richter. If there is a month that is more intense than others. :::

Null hypothesis: there is no difference in between each month. The month has no effect if the intensity richter is more or less severe.

Alternative hypothesis: there is a difference in between each month. The month has an effect on if the intensity richter is more or less severe.

```
ggplot(csv_table, mapping = aes(x = month, y = intensity_richter)) +
  geom_boxplot(outlier.color = "red")
```

