

# Sprawozdanie nr 1

Klaudia Janicka 262268 i Natalia Iwańska 262270

2022-12-13

## 1. Wstęp

Niniejszy raport stanowi analizę danych rzeczywistych dotyczących trzęsień ziemi o skali wydarzenia powyżej 6 na przestrzeni lat 1900-2013. Dane pochodzą ze strony [kaggle.com](https://www.kaggle.com). Zgodnie z informacjami zawartymi na podanej stronie źródłem danych jest United States Geological Survey.

Celem naszej analizy jest odpowiedzenie na pytanie jakie regiony są najbardziej narażone oraz zbadanie jakie czynniki mają wpływ na występowanie trzęsień ziemi.

### 1.1 Opis zmiennych

Do analizy wykorzystujemy dane zawarte w kolumnach:

- *place* - tekstowy opis regionu geograficznego w pobliżu zdarzenia;
- *latitude* - szerokość geograficzna podana w stopniach, przyjmuje wartości z przedziału  $[-90, 90]$ , gdzie wartości ujemne oznaczają południowe szerokości;
- *longitude* - długość geograficzna podana w stopniach, przyjmuje wartości z przedziału  $[-180, 180]$ , gdzie wartości ujemne dotyczą zachodnich długości;
- *mag* - skala zdarzenia;
- *nst* - liczba stacji sejsmicznych użytych do określenia lokalizacji trzęsienia;
- *time* - czas wystąpienia trzęsienia;
- *depth* - głębokość zdarzenia w kilometrach;
- *date* - data zajścia zdarzenia;

gdzie jako zdarzenie rozumiemy wystąpienie trzęsienia ziemi.

Jako zmienne kateryczne uznajemy zmienną *place*, która po transformacji wskazuje na kraj, w którym zdarzenie miało miejsce oraz *date*, która po transformacji oznacza miesiąc zajścia zdarzenia. Pozostałe zmienne to zmienne ciągłe. (Tak myślę, ale może być inaczej)

### 1.2 Obsługa błędów

Wiersze z brakami danych usuwamy za pomocą funkcji `drop_na()` z biblioteki `dplyr`.

## 2. Analiza danych

### 2.1 Miejsca najbardziej narażone na trzęsienia ziemi

W celu ustalenia państw najbardziej narażonych na wystąpienie zdarzenia posłużymy się analizą graficzną.

Mapa świata z naniesionymi miejscami trzęsień ziemi

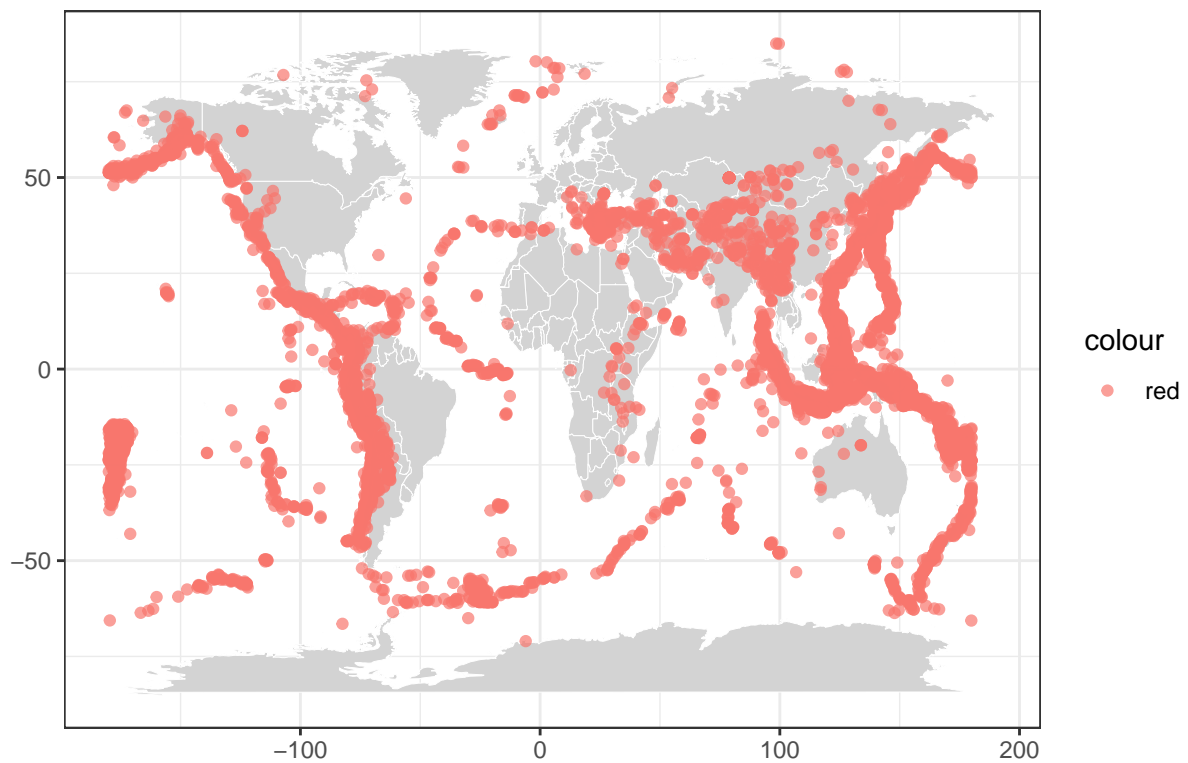
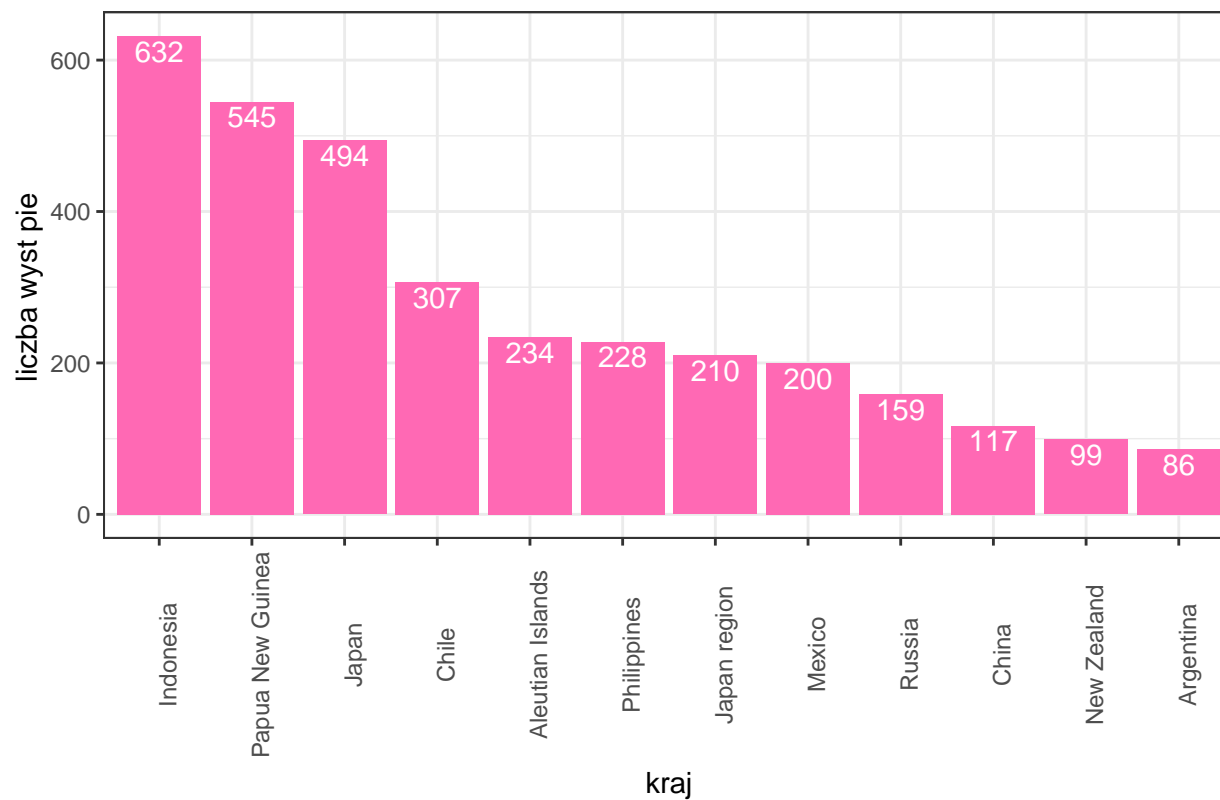


Figure 1: Mapa świata z naniesionymi miejscami trzęsień ziemi

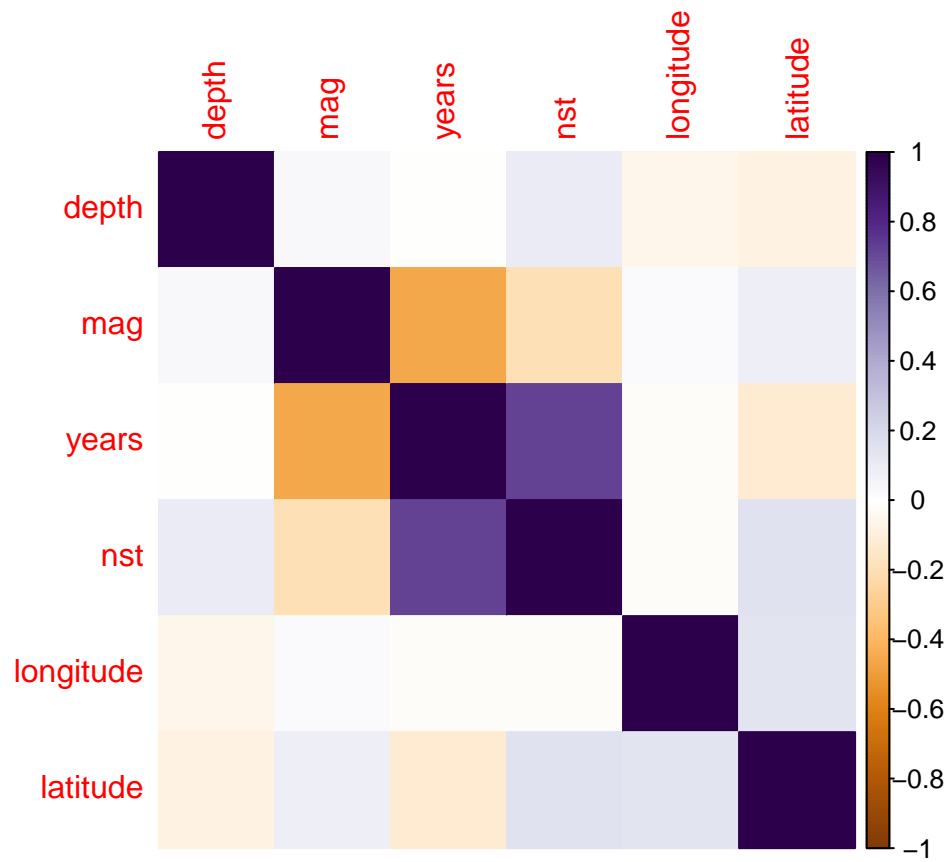
Z powyższego wykresu i posiadanej wiedzy z zakresu geografii jesteśmy w stanie stwierdzić, że najbardziej narażone są między innymi Chile, Japonia i inne kraje azjatyckie.

Wykres słupkowy dla 12 najbardziej narażonych miejsc



Wykres ?? potwierdza nasze wcześniejsze przypuszczenia oparte o wykres 1.

## 2.2 Macierz korelacji



## 3. Wnioski