

USGS Earthquake Data Analysis

Akshay Yadav

April 30, 2017

1 Abstract

This project is an attempt to build a R shiny application that downloads earthquake data from USGS database between user defined dates, processes the data into a convenient dataframe and displays using different types of visualizations. The earthquake records are displayed in histograms that show country and continent-wise distributions. The exact continent-wise locations of the earthquakes can also be viewed on an interactive world map. Two scatter plots in a different tab help the explore relationship between the significance, depth and magnitude variables of the earthquakes. A slider widget is also provided which helps to filter-out less significant earthquakes for analysis.

2 Introduction

The United States Geological Survey (USGS) is an agency that maintains records of the earthquakes that are detected in the US and around the world. These records are updated on a daily basis and can be accessed easily. Past and present records between any dates can be obtained through a webAPI service in JSON (geoJSON) or other convenient formats. In this project, I have attempted to build a R shiny application and related methods that download earthquake records between any specified dates and provide different visual representations of the records after processing the data. The following sections will describe different options and back-end data processing in stepwise manner

3 Data - Accessing, Format and Processing

The earthquake records from USGS database can be accessed using a customized webAPI URL which returns the earthquake data between specified dates. The user can select the dates using dateRange selector widget which is displayed on application initialization. The "Get Earthquake Data" button, when clicked after selecting the desired dates, will initialize the data download and processing on the server-side. After the button is clicked, the custom URL to accommodate the user selected dates is prepared and the earthquake records are

downloaded from the agency database in the JSON format. The JSON contains information on different types information on earthquakes recorded between the specified dates in form of different qualitative and quantitative features. The JSON is processed using functions from the "jsonlite" package to obtain a list of three dataframes viz. feature, geometry and id. All the dataframe are column bound to obtain a single dataframe.

An additional processing step included in this project was to infer the names of the countries and continents for each earthquake in the data using the latitude and longitude coordinates in the feature dataframe. The "rworldmap" and "sp" packages were used for this purpose. The coordinate ranges for each country and continent were obtained using functions from "rworldmap" package and the actual earthquake coordinates were mapped to these coordinate ranges using functions from the spacial "sp" packages. Two new columns specifying the country and continent names for the earthquakes were added to the earthquake records dataframe obtained earlier to get the final processed dataframe. The earthquakes that has occured in oceans and disputed territories were marked as "Unclassified" in the countries and continent columns.

4 Data Visualization

When the data processing is complete a histogram that shows a country-wise distribution of earthquakes is displayed automatically along with a radioButton widget below the "Get Earthquake Data" button. The radioButton widget has two options and can be used to change the histogram to show continent-wise distribution of the earthquakes. The histograms are interactive and can be used to show/hide desired countries or continents in the histograms.

One of the most important quantative features for each earthquake record is the "significance" value of the earthquakes. This value is assigned to the earthquakes by USGS and is function of other earthquake variables such as magnitude, maximum MMI, number of felt reports and estimated impact. The significance value therefore can tell us how significant the earthquake was in a more unbiased sense. In this application, the user can used a slider widget displayed below the radioButton to filter out less significant earthquakes. The default value of the slider is set to zero so the no records are filtered out. The effect of slider widget is automatically reflected in country-wise and continent-wise distribution histograms.

The exact location of earthquakes, on a world map, can be viewed using the drop-down and "Show on Map" widgets displayed below the slider widget. The user can select the desired continent from the drop down and clicking the "Show on Map" button will display the earthquakes related to the selected continent on an interactive world map. Selecting a different continent and clicking on the button again will update the map to show the earthquakes related to newly selected continent. The effect of the slider is also reflected automatically in the map.

In a second "Scatter Plot" tab, two interactive scatter plots that display

relationship between the significance v/s depth values and significance v/s the magnitude values are displayed. The points are colored according to country or continent names (according to selection in the radioButton widget) and can be modified to show/hide points related to one or more countries or continents. Here too, the effect of using the significance slider widget is automatically reflected in both the scatter plots.