COMPSCI 2XB3:Computer Science Practice and Experience: Binding Theory to Practice
Project Proposal

Project Title:	*Earthquake Risk Assessment*
Lab Section Number:	**L01**
Student Names:	*Kan Hailan, JiHwan Kim, Sembakutti Kalindu, Tao Haoyang, Ye Fang*
Group Number:	**01**
Student McMaster Emails:	*kanh3@mcmaster.ca, kimj162@mcmaster.ca, sembukkk@mcmaster.ca, taoh4@mcmaster.ca,yef10@mcmaster.ca*

By virtue of submitting this document I electronically sign and date that the work being submitted is my own individual work.

Abstract
Earthquakes have always been a major challenge for human beings because of the deaths and economic loss they cause. Our project is an app that helps individuals and governments analyse earthquake risk to minimize loss due to earthquakes. It uses past earthquake information along with population density from Canada government open dataset to calculate the risk. It allows the user to enter a location and select a radius, then it lists historical occurrences of earthquakes in that range and gives a score of 0 to 9 from low to high for the risk based on distances between that location and earthquake centers, magnitude and frequency of past earthquakes in that range, and population density of which province it is in.

1. Objective

Our objective is to develop a software application that provides historical earthquake information searching and risk rating service for individuals, commercial companies and government.

2. Motivation

Earthquakes cause tremendous deaths and economic losses every year around the world. In the past fifty years, 1.5 million people died from earthquakes and the estimated number of injuries is 4.5 million according to the investigated average ratio of mortality and morbidity in earthquake disaster. The recent big earthquake caused around 230,000 deaths, 3000,000 injuries and as much as $13.9 billion damage in Haiti in 2010. Although the recent big earthquake in Canada caused limited damage due to the very sparse population, the earthquake risk is still big in future due to the rapid growing population. As a natural disaster, we cannot stop the occurrence of earthquakes. However, we can minimize the loss and damage as much as possible by taking some steps in advance, such as reducing the population density and raising the earthquake-proof performance level for the buildings, by foreknowing the earthquake risk. A practical application which provides earthquake risk information will be beneficial for the communities. Based on the current dataset, our product will face individuals and groups, such as commercial companies and governments, who have the demand for foreknowing the earthquake risk in Canada.

Our app mainly has searching and risk rating functionalities, and it is very easy to use. Users just need to enter the information in the corresponding categories, such as city name, geographical location, date, radius, the app will show users the detailed historical earthquake information and risk rating information.

3. Prior Work

Temblor

Temblor is a mobile app that lets users enter an address, and then find out the earthquake risk for that specific location. The app is specifically designed for the public to determine the level of earthquake risk to their homes, work location or children's educational institutions. Although, this app takes into account fracture lines, earthquake-induced landslides, and quality, cost and size of the building, it doesn't consider external factors that overwhelm the earthquake response teams and hinder any assistance after the fact. In a similar manner, our application will provide an earthquake rating, but also incorporate population density, and availability of emergency response teams to calculate an earthquake rating. In addition, our app will also use historical earthquake occurrences around a geographical location to calculate risk level to that specific location.

ShakeMap by USGS (U.S. Geological Survey)

This web application provides historical occurrences of earthquakes depicted as circles of variable sizes based on their intensity level. In a similar manner, our app will list historical occurrences that had occuranced within a specific radius of the user's geographical location sorted based on the proximity. Our app will not provide a graphical interface, but a simple list of the historical occurrences. Our search user interface functionality will be limited to selecting a specific location and radius by the user. But, our output list will provide the intensity by highlighting each list item in a specific colour that signifies the intensity.

4. Input/output and proposed solutions

Inputs:
Earthquakes in Canada
https://open.canada.ca/data/en/dataset/4cedd37e-0023-41fe-8eff-bea45385e469

This dataset provides historical earthquake data dating back to 1985, and using this dataset we will provide the users the ability to search the earthquakes that had happened in close proximity to their geographical location.

Population and dwelling counts, for population centres, 2016 and 2011 censuses
https://open.canada.ca/data/en/dataset/7f03d494-8591-45cc-ac1b-4a99237d35b0
The same population and dwelling counts, plus data for subdivisions (municipalities)
https://open.canada.ca/data/en/dataset/402495f0-6415-4fd8-bcfc-25a304e6fc8b

We will obtain the population density from this dataset to calculate an earthquake risk rating that is based on the overloading effect of after-the-fact emergency response teams.

Outputs:
A list of historical earthquake occurrences within a given radius from a geographical location determined from a geographic name input.
An earthquake risk rating calculated based on close proximity to previously occurred earthquakes, and an overloading factor of emergency response teams due to a large population density.

Brief description of the proposed solution in terms of how the inputs will be used to generate the outputs:

Our Canadian earthquake dataset provides records of historical occurrences of earthquakes near Canadian geographical locations. The data set provides latitude and longitude of an earthquake. For the same earthquake, the data set also provides an approximate distance from a geographical name such as Victoria, Gold River, or White Rock. We will compare the geographical names given in this dataset with that of the same given in the population density dataset. As a result of this comparison, we will be able to model an earthquake risk rating for populations that are in close proximity to the locations of historical earthquake occurrences. The earthquake risk rating will be a score calculated between 0 and 9, with 0 being the lowest and 9 being the highest risk.

In addition, the users of our software application are able to search historical earthquake occurrences within their place of residence, work location or children's schools. We will first filter the earthquakes dataset by the geographical names of each record. For instance, when the user enters a location in their input, we are able to use a map function to filter the historical earthquakes that had happened in close proximity to the user's location. We will traverse through the filtered list to find the geodesic distance between the user's location and each of the earthquake locations. Finally, we will display the filtered list of earthquakes sorted based on the proximity to the user's location. These historical earthquake data will be displayed as a list of earthquakes with colours identifying their intensities.

5. Algorithmic challenges:
Our earthquake dataset can be sorted in many different ways, and contains date, magnitude and city name.etc. And since users can apply multiple filters to the data such as the magnitude and a time range, the algorithm challenge will be try not making our time complexity linear. It's better to filter the data to a smaller subset first, and then utilize sorting and searching algorithms on the smaller subset just for the best performance. By applying merge sort, we can get a complexity of O(n log n) for sorting the data in general. For subarrays of the merge sort, using less complex algorithms are preferred(insertion sort.etc.). For searching algorithms, binary search is our first choice since it has the least complexity on a sorted dataset.

Also, we are combining two dataset: Earthquake and Canada population density. Our Earthquake dataset contains city names for an earthquake, but not every one corresponds to a province (there are gaps in the province column). As a result, we consider them (provinces in Earthquake dataset) unusable and will discard all the province data in Earthquake dataset. Instead, we will implement our own function to map a given city name to a province. Our preliminary discussion points out a solution by using HashMap. After that, we can union our population dataset to our earthquake dataset.

6. Project plan

Milestone	Deliverable	Date
Set up: data set, Helper Modules	Find out data sets; Earthquake, Population Helper Modules: To extract data from the excel file and transform them into instances of ADT	Feb 16th 2020
Requirements specification document	Document about the requirements specification according to the CS 2ME3 instruction	Feb 20th 2020
ADT Implementation	ADT implementation for data to be handled; ADT will be implemented with array	Fed 23th 2020
ADT Implementation continued, Merge sort Algorithm	ADT implementation finished, implementation of merge sort Algorithm; sorted by City	Mar 1st 2020
Algorithm: Union operation	Implementation of Union operation of two data sets	Mar 8th 2020
Searching algorithm: binary search	Implementation of binary search for sorting purpose	Mar 15th 2020
Analysis of algorithm	Analyze the time complexity; O(f(n))	Mar 22th 2020
Unit testing : Junit	Construct test cases using Junit	Mar 22th 2020
Report	Prepare for presentation and design document	Mar 29th 2020

.

References

[1]Deaths Due To Earthquakes Worldwide 2000-2015. Retrieved February 7, 2020, from https://www.statista.com/statistics/263108/global-death-toll-due-to-earthquakes-since-2000/
[2]Earthquakes: Information & Facts - Canadian Red Cross. Retrieved February 7, 2020, from

https://www.redcross.ca/how-we-help/emergencies-and-disasters-in-canada/types-of-emergencies/earthquakes/earthquakes-information-facts
[3]Haiti Earthquake Of 2010. Retrieved February 7, 2020, from https://storymaps.arcgis.com/stories/6b2764350f2f4fd08af585790785c86a
[4]Temblor App - New Version released on July 24, 2019. Retrieved February 7, 2020, from https://temblor.net/temblor-app/
[5]Recorded Ground-motion Parameters - ShakeMap Documentation. Retrieved February 7, 2020, from http://usgs.github.io/shakemap/manual3_5/tg_parameters.html#data-acquisition
[6] Schürger, T. (2019, July 18). How does containsValue() method of HashMap work in Java? Retrieved February 6, 2020, from https://www.quora.com/How-does-code-containsValue-code-method-of-code-HashMap-code-work-in-Java