Are We Ready For the Big One?
A Study on Earthquake Preparedness in the Pacific Northwest

Research Proposal

4/11/2021

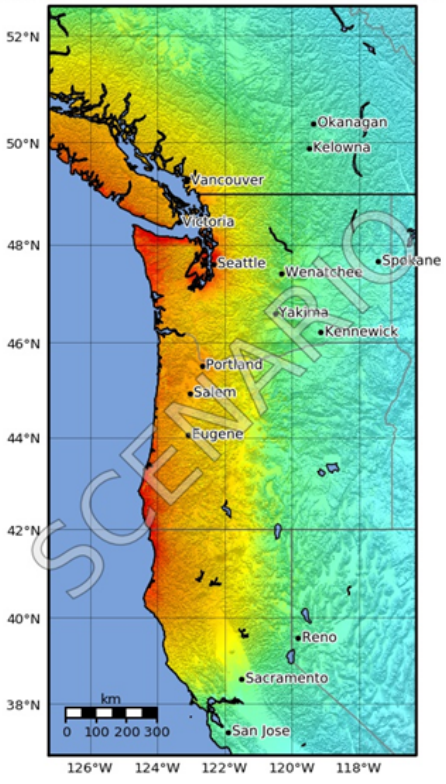
Brian Truong and Maria DiMedio

Overview

This research proposal is intended for the sales and marketing departments of a construction company providing earthquake retrofitting services interested in driving the ad/awareness campaign “Are we ready for the big one?”. The campaign focuses on the real possibility of a devastating earthquake striking the Pacific Northwest, originating from the Cascadia Subduction Zone, just off the coast of the region [1]. This earthquake is estimated to have a strength of more than 7.1 magnitude on the Richter scale, with a 37% chance of occurring in the next 50 years [1] [2]. Due to more than three centuries since the last earthquake, a large number of buildings are not built to earthquake standards [1]. As such, the objective of this study is then to understand exactly which counties in the region are most susceptible to such an earthquake, accomplishing this through data collected from building seismic code compliance and survey data. From this analysis, the sales and marketing departments will know which counties in the Pacific Northwest should be targeted for direct ads, increasing public awareness, and which in turn increase customers.

SHAKING	Not felt	Weak	Light	Moderate	Strong	Very strong	Severe	Violent	Extreme
DAMAGE	None	None	None	Very light	Light	Moderate	Moderate/heavy	Heavy	Very heavy
PGA(%g)	<0.0066	0.0795	0.954	4.99	8.76	15.4	27	47.4	>83.2
PGV(cm/s)	<0.0028	0.0383	0.524	3.03	6.48	13.9	29.6	63.4	>136
INTENSITY	I	II-III	IV	V	VI	VII	VIII	IX	X+

Scale based on Atkinson and Kaka (2007) Version 1: Processed 2020-11-15T02:19:47Z
△ Seismic Instrument ○ Reported Intensity ★ Epicenter



(Right, top) Cascadia Subduction Zone Area of Impact [2]

Research Question

Which counties in the Pacific Northwest are least prepared for a potential earthquake, according to the proportion of buildings meeting seismic code requirements, and data collected from surveying local citizens?

In order to answer this question, a few key terms should be defined. ‘Preparedness’ in this study will be assessed in two ways. First, it is defined as the readiness of a building to contend with an earthquake, according to seismic construction codes, including retrofittings. Second, preparedness is defined as citizens’ awareness

of the potential for an earthquake, the number of days of self sufficiency (the government recommendation is at minimum 2 weeks of self-sufficiency [1] [3]), and if plans for post-disaster action exist.

The Pacific Northwest in this study refers to the 88 total counties in Oregon, Washington, and Northern California. The 13 counties in Northern California are: Del Norte, Siskiyou, Modoc, Humboldt, Trinity, Shasta, Lassen, Tehema, Plumas, Mendocino, Glenn, Butte, Lake. The counties selected are based on the estimated impact area of the Cascadia earthquake [2].

Data

For building resilience, the main data source to be leveraged will be the publicly available building permit data. This data contains information about:

- The date a building was built.
- Address where the building is located.
- Any major upgrades to a building.
- Building safety codes (including seismic codes)

This will then be used in tandem with historic building code records to find the earthquake resilience of buildings [4] [5]. The data can be accessed online or by a request to state/local government [6]. However, some of this data is not available in an easily digestible format, and so might need to be processed by hand, or received in person from local governments.

For data on the general preparedness of the population, surveys will be conducted to collect additional data on this population. Surveys will be distributed by mail, in multiple languages, with phone follow ups if possible. If there are limitations on the sample size of the mailing/phoning campaign, addresses will be chosen based on a random selection process. Surveys distributed by mail will be directed to addresses pulled from permit data and public address databases. Phone follow ups will be based on the same address data but pulling phone numbers from a database that connects addresses to phone numbers, for instance, White Pages. For this data, a bias can arise from non-respondents if they represent a segment of the population. To mitigate this, efforts will be made to follow up on non-responders, with phone or other means. All survey participants will be informed of the nature and objective of the survey, as well as the commitment of the study to ensure that their individual data will be anonymized. Given that the data will be compiled into a county-by-county summary, individual respondents will be effectively anonymized.

Study Design

The proposed study is an explanatory sequential mixed-method research design consisting of two phases. This will be an observational study, in which we will collect data on the current state of building codes and citizens' survey responses in two phases. The first phase is focused on quantitative data collection. In this phase, we will collect data on building codes for buildings built in the Pacific Northwest. This data will be collected through public access to building

permits, both residential and commercial. We will collect the most recent data available on building codes, specifically looking at the seismic codes indicating whether a building is outfitted with additional safety measures for earthquakes. We will group the building data by county and perform a county-by-county analysis of the proportion of buildings that meet these standards.

The outcome of this analysis will be the scoring of counties in the Pacific Northwest by the proportion of buildings meeting seismic codes overall. Based on these calculations, we will analyze all 88 counties and identify which have a preparedness calculation (proportion of buildings meeting seismic codes) that are less than the group median. We will select any of the counties rated below the median to be part of the second phase of our study. These counties, by our definition, will have the greatest need for retrofitting services.

The second phase of this study will be qualitative data collection through a survey of counties identified in Phase 1, collecting response data on citizens about their preparedness for an earthquake, as well as their existing knowledge of the threat to the Cascadia Subduction Zone. Using mailing and phone address data from publicly available building records, we will send surveys to building owners to collect data on their existing knowledge of the risk of the earthquake hitting, and their preparedness for an earthquake. We will send all addresses a physically mailed questionnaire, with a stamped return envelope. We will then collect the returned questionnaires and track non-response and will follow up with non-responders with a phone call to verbally collect survey responses. At the end of the study, we will assess the number of non-responders in each county and determine the proportion of total buildings with non-response and include this in our results.

Sample

Our sampling frame will be commercial and residential buildings located in the Pacific Northwest. We expect this to be a very large population for the first phase of our study, and expect for the initial phase to take several months to collect millions of building records and compile seismic code data. Because this data is publicly available and free, we would recommend allocating additional resources to this collection process. Our data science team will develop an automated record processing program that can alleviate the manual records checking process, increasing the feasibility of this large project. In acknowledgement of the scale of this study and with budget in mind, as a secondary option, we will select buildings using stratified random sampling of residential and commercial buildings in each county. We will assess the necessary sample size by a statistical power analysis based on existing literature.

The second phase of the study, the survey, will have a sampling frame representing the counties selected in the first phase. This will be roughly half the counties in phase 1. Letters with our survey will be sent to every building owner listed. Follow up phone calls will only be made to non-responders.

Variables and/or Intervention

The variable of interest in this study is preparedness for an earthquake. We have defined that in two categories. First, building preparedness (whether or not the building is up to seismic codes), and second, citizen preparedness (captured through surveys). The building's preparedness will be determined in a categorical variable, indicating it meets or does not meet seismic codes. The county's preparedness will be the proportion of buildings which meet seismic codes.

The second category of preparedness will be evaluated through qualitative response to survey results. The survey will be given asking the following questions:

a. Are you aware of the risk present in the Cascadia Subduction Zone for a potential future earthquake?
b. Is your building located in proximity to emergency services or a hospital?
c. Does your building have food and water stored, enough to support 2 weeks of self sufficiency?
d. Do you believe the possibility of an earthquake hitting to be a real threat?

By assessing this variable in these ways, our study will provide insight for the marketing team on where to focus its marketing campaign, and the level of citizens' awareness of the risk for a potentially devastating earthquake.

Statistical Methods

This study will use descriptive statistics and standard statistical methods to analyze the data, as well as accounting for potential biases in the sample selection, such as those posed by missing building data or non-respondents. Any missing data interpolation will leverage methods discussed in the paper "Predicting residential building age from map data" [7]. Our statistical power analysis will be informed by the most recent literature in this field. In general, our statistical methodology for the analysis of the survey data will follow best practices set in the Creswell & Creswell textbook [8].

Finally, we are aware we will likely not reach a 100% response rate in our survey collection. To handle this, we will follow up in person and via phone with non-responders, as well as review the data for non-responders to determine if there are any biases to account for in our final analysis.

Potential Risks

A source of potential risk is the completeness of the building permit data, as data may not account for all buildings in the region. For instance, buildings may have been built before a permit system existed, without a permit, or even with a permit, but with records for such a permit missing. A remedy for this shortcoming is to reinforce the permit data with survey data, selecting a sample of addresses that are missing. This is especially the case if data is missing on relevant

segments of buildings, such as those made before earthquake building codes were enforced. Additionally, known building data can be used to fill in missing data using some algorithm.

Another potential risk lies in the process of obtaining the building permit data. While lots of building permit data can be found online, it is often more recent permit data, and as such covers mostly earthquake resilient buildings. Thus, to find older building permits, study team members may have to reach out to different jurisdictions in person to obtain the data. In some cases, this may require submitting a Freedom of Information Request (FOIA) [6], which can take 20 days to process [9]. If records are missing entirely from government records, the study team may need to reach out to construction companies. Thus, to mitigate possible delays, it is necessary to assess which jurisdictions will require FOIA requests, and initiate them early in the process, and plan out any trips to local jurisdictions to efficiently collect the building permit data. If it becomes too difficult to obtain certain records, it might be necessary to rely on the previously discussed methods to fill in the gaps.

Finally, there is also a potential risk for bias in this study by the non-responders of our survey, as well as between surveys completed via mail or via phone/verbally. To combat any potential bias in our results, we will assess survey feedback and non-responder data and look for any trends in response between these subgroups. We will also select a sample of non-responders to survey in person, to determine any additional bias.

Deliverables

The anticipated timeline for this study is 6 months in total. The first phase of the study is estimated to take 2-3 months to complete, given the large sample size, the potential for physically accessing and compiling data, and the possibility of filing FOIA requests. The second phase is estimated to take another 2-3 months, to allow for mailing and returning surveys, and following up with non-responders.

The first deliverable to be expected is the ranking of counties identified in phase 1. This will be reported to executives in the sales & marketing department, along with a summary of our findings in phase 1. This summary will report the proportion of buildings up to seismic codes in all counties, and will identify which counties are ranked below the median and included in phase 2. The final deliverable will be a report on the current state of preparedness of citizens in those counties for a potential earthquake. This will include data on how ‘prepared’ these citizens are for an earthquake, as well as summary statistics on how many citizens deem an earthquake in this region to be a potential threat. We will also identify how many citizens are aware of the potential threat to the Cascadia Subduction Zone, to be used in the ‘Are We Ready for the Big One’ ad/awareness campaign.

Statements of Contribution

- Brian Truong:
    - Primary section contributions:
        - Overview, Data, Potential Risks
    - Also contributed corresponding slides
    - Helped edit and provide feedback for other sections of the report
    - Provided research / sources for different aspects of the report
    - Together worked on the research question and statistical methods
- Maria DiMedio:
    - For this report my primary section contributions were the study design, sample, variables, and deliverables sections. I also completed the corresponding slides for the presentation. There was a joint effort for the research question and statistical methods sections, as well as background information and editing of other sections of the report.

Works Referenced

- [1] Oregon Office of Emergency Management : Cascadia Subduction Zone
- [2] M 9.0 Scenario Earthquake - M9.0 Cascadia (geometric mean) (usgs.gov)
- [3] Preparedness | Washington State Military Department
- [4] Legislation - Western States Seismic Policy Council (wsspc.org)
- [5] Earthquake Building Codes | FEMA.gov
- [6] Building Permit Records: How to Find and Use Permit Records
- [7] Predicting residential building age from map data - ScienceDirect
- [8] Creswell, J.W. and Creswell, J.D. (2018) Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage, Los Angeles.
- [9] Responding to Requests (justice.gov)