Earthquake Simulation: Effect of Floor Number on Collapse Time

6/1/17

ASA Statistics Competition

Table of Contents

Abstract

Earthquakes are a major potential source of danger to buildings and people around the world. The amount of stories (floors) in a building can be a predictor of how quickly the earthquake affects its structure, and how quickly the building collapses. Earthquakes vary in time interval, so it is useful to know how the amount of stories in a building could affect how long it takes for a building to collapse. We modeled earthquakes using a Jenga tower of stories 2, 3, and 4. We used an EME molecular demonstrator in order to shake the box that the tower was on, and we fit the data to a transformed 1/y regression model. Additionally, we ran one factor ANOVA and 2-sample difference of means t-tests on our sample data. We were able to extrapolate the time interval until collapse to more stories using confidence intervals of a predicted value that we got from our transformed regression. We found that as the floors of a building increases, the time interval it takes for an equal magnitude shake to cause it to collapse decreases. Therefore, in order to preserve the structural integrity of buildings, we must stick to lower floor levels.

Introduction

Earthquakes can hit any time. Humans go around in buildings virtually all the time, even though these buildings are extremely susceptible to the damage that even a quaint earthquake can cause them. The rapid shaking of an earthquake can easily cause collapse in a building, which can cost a lot of money and harm many people who are inside.

Therefore, we came up with our problem: how does the number of floors in a Jenga tower affect the time interval it takes for the tower to collapse (keeping the magnitude of the shake and tower shape constant). Our experimental units were the actual tower. We manipulated the number of floors on the tower (and therefore the mass of the tower as well as the frequency of the shake). Some potential extraneous variables that we could not control was external noise, humidity, time of day at time of test, difference in individual Jenga blocks, and inaccuracies in timing. However, we hoped to resolve or at least limit the effects of these through random assignment of treatments. We were able to keep constant the amount of blocks per floor, the shape of the tower as floor increased, and the magnitude of the shake (the frequency of the shake varies because frequency changes with mass).

According to the University of Portland, taller buildings are more susceptible to lower frequency but longer lasting earthquakes, while buildings with fewer floors are affected more easily by high intensity but short duration earthquakes. ("Buildings")

Studies on height of buildings on earthquakes showed relatively similar results for 2 and 3 story buildings; in a case study of a 1964 earthquake in Niigata City, they showed how damage sustained to these two types of buildings was very similar. There were lots of fissures in the ground as well as physical damage on the building, such as fractures. (Kishida)

We chose to use Jenga blocks to simulate the earthquake because they were inexpensive, portable, easy to assemble, and consistent in size, shape, and mass. The use of Jenga blocks allowed us to quickly arrange towers of different heights after they were knocked over in order to efficiently get enough data for this experiment in the allotted time we had.

We used an EME Molecular Motion Demonstrator because it was a reliable way to shake the tower. The machine oscillated a metal square diagonally back and forth. We were able to adjust the magnitude of the shake (frequency) and make sure that the magnitude of the shake was consistent in our trials.

Although other studies and research have shown that taller builders and shorter buildings can sustain different loads, or at least react somewhat similarly, we believed that the time interval until the collapse of the tower would decrease when we increased the number of floors. This is because it appeared that our building would become increasingly unstable as we added more floors, and would therefore lead to lesser structural integrity, vastly decreasing the time interval it takes for the building to collapse. Additionally, we thought that this decrease would be significantly significant between all different treatments (2 floors, 3 floors, and 4 floors).

Materials and Methods

Materials

Jenga blocks (3 per floor), scale, holding container (plastic), EME Molecular Motion Demonstrator, stopwatch, random number generator, duct tape

Procedure

1) Gather all materials
2) Tape three Jenga blocks to bottom of holding container flush and parallel to each other directly in the middle
3) Tape holding container to EME molecular motion demonstrator
4) Use random number generator to randomly assign order of numbers 1-3
5) Place number of floors on top of taped down Jenga floor based on first number in the randomly assigned order. If order is 2 1 3, place 2 floors on top first (total of 3 floors).
   - A. For 1 additional floor: Place three Jenga blocks parallel and flush to each other in perpendicular direction on top of already placed Jenga blocks.
   - B. For 2 additional floors: Repeat A two times.
   - C. For 3 additional floors: Repeat A three times.
6) With additional floors added onto taped Jenga floor, turn EME Molecular Motion Demonstrator on and start stopwatch. Make sure to use the same shaking amount on the machine each time by carefully moving knob on the machine.
7) Stop the stopwatch when one block hits the ground on the holding container. Record time interval. Remove blocks that had been added.
8) Repeat steps 5-7, but add the number of floors based on the second number in the randomly assigned order of numbers 1-3 instead of the first number.
9) Repeat steps 5-7, but add the number of floors based on the last number in the randomly assigned order of numbers 1-3.
10) Repeat steps 4-9, 29 more times for a total of 30 trials.
11) To record frequency, first place 1 additional floor into holding container (for a total of 2 floors). Turn EME on and count how many oscillations the device makes in a total of 10 seconds. Then divide this number by 10 and record.
12) Repeat step 11 two more times for a total of three trials. Calculate the average frequency.
13) Repeat steps 11-12 for 3 floors and 4 floors.
14) To find mass, place holding container with 4 floors on it on scale. Repeat three times, record mass each time, then calculate the average.
15) Repeat step 14 with 3 floors and 2 floors to find masses of these.
16) Analyze data.

Results and Analysis

Tests

Floor 2 vs Floor 3 vs Floor 4 Time Interval

One-factor ANOVA Test

H_0: $\mu_1 = \mu_2 = \mu_3$ H_a: At least one μ_i is different than the other ones.

We will use an α level of 0.05

Checks: Our n>30 for all our floor levels of data. Therefore, by central limit theorem we can assume that the sample data is normally distributed. Additionally, our observations are independent of each other. We also randomly assigned treatments. The standard deviations are not approximately to each other ($s_1 = 0.8444$, $s_2 = 0.6572$, $s_3 = 0.2159$), however, we must proceed in our test anyways. We successfully pass three out of four checks.

$$F = \frac{\frac{SSTR}{df_r}}{\frac{SSE}{df_e}} \; with \; df_r = k - 1 \; and \; df_e = (n_1 + n_2 + n_3) - k$$

Results: F = 9.0149, df = 2 & 87, p-value = 0.0002

Our p value of 0.0002 is less than our alpha value of 0.05. Therefore, we reject the null hypothesis. We can assume that the true population time interval for the tower to fall down is not the same for 2, 3, and 4 floors.

Floor 2 vs Floor 3 Time Interval

2 Sample Difference of Means t-Test

$H_{0:}$ μ_1-μ_2=0 H_a: μ_1-μ_2≠0

We will use an α level of 0.05.

Checks: For floor 2 and floor 3, our n (30) >24. Therefore, by central limit theorem, we can assume that the sample data is normally distributed. Additionally, we randomly assigned treatments. It is safe to proceed.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}} \; with \; df = n_1 + n_2 - 2$$

Results: t = 2.622, df = 58, p-value = 0.0111

Our p value of 0.0111 is less than our alpha value of 0.05. Therefore we reject the null hypothesis. We can assume that the true mean time interval for the tower to fall down with 2 and 3 floors is not equal to each other.

Floor 3 vs Floor 4 Time Interval

2 Sample Difference of Means t-Test

$H_{0:}$ μ_1-μ_2=0 H_a: μ_1-μ_2≠0

We will use an α level of 0.05.

Checks: For floor 3 and floor 4, our n (30) >24. Therefore, by central limit theorem, we can assume that the sample data is normally distributed. Additionally, we randomly assigned treatments. It is safe to proceed.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}} \; with \; df = n_1 + n_2 - 2$$

Results: t =1.1508, df =58, p-value = 0.2546

Our p value of 0.2546 is greater than our alpha value of 0.05. We cannot reject the null hypothesis. We cannot assume that the true mean time intervals for the tower to fall down for floors 3 and 4 differs.

Floor 2 vs Floor 4 Time Interval

2 Sample Difference of Means t-Test

H_0: μ_1-μ_2=0 H_a: μ_1-μ_2≠0
We will use an α level of 0.05.

Checks: For floor 2 and floor 4, our n (30) >24. Therefore, by central limit theorem, we can assume that the sample data is normally distributed. Additionally, we randomly assigned treatments. It is safe to proceed.

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}} \; with\; df = n_1 + n_2 - 2$$

Results: t =4.1329, df = 58, p-value = 0.0001

Our p value of 0.0001 is less than our alpha value of 0.05. Therefore we reject the null hypothesis. We can assume that the true mean time interval for the tower to fall down with 2 and 4 floors is not equal to each other.

Regression Analysis

Our manipulated variable, floor number, was an actual number (2, 3, and 4). Therefore, we were able to fit a regression line that predicted the time interval for the tower to fall based on the number of floors that the tower had.

We first generated this linear model and calculated predicted time intervals of collapse for additional floor numbers: $\hat{y}=2.5455-0.3288x$

Predicted Values Based on Basic Linear Model

Floor Number	Predicted Value
5	0.9015
6	0.5727
7	0.2439
8	-0.0849

We encountered problems with this model. Although the slope was significant as proven by the model utility test using an alpha of 0.05, ($p = 0.00012$), there was a very low r^2 value of 0.1555. This means that only 15.55 % percent of variability within the time interval for collapse could be explained by a linear relationship with the number of floors in the building. Additionally, the r value of 0.3943 indicated only a weak correlation between number of floors and time interval until collapse. The standard error value of 0.6327 was very large relative to the data collected, as almost all of the recorded time intervals were under 2 seconds. Finally, when we attempted to extrapolate the model, we found that using this simple linear model did not work the best; when we extrapolated to 8 floors, the model gave us a predicted time interval of a negative value, which is impossible as the time interval cannot be negative, and should instead only approach 0 seconds.

Therefore, in order to find a more accurate model that we will not go into a negative predicted time interval, we apply log and reciprocal transformations to find the best transformed model.

Transformation	S_e Value	R^2 Value	|r-value|	p-value (Slope)	Slope Significant?
None	0.6327	0.1555	0.3943	0.0001	Yes
Log(y)	0.3006	0.1904	0.4363	1.71e-05	Yes
Log(x)	0.6296	0.1639	0.4048	7.55e-05	Yes
Log(x) & Log(y)	0.2983	0.2026	0.4501	8.52e-06	Yes
1/y	0.1818	0.187	0.4324	2.08e-05	Yes
1/x	0.6846	None	None	None	None
1/x & 1/y	0.2005	None	None	None	None

Based on our transformations, the best model to use is a transformed model where we use 1/y instead of y. Compared to the other transformations, as well as to the original simple linear model, the s_e is extremely low at only 0.1818. This s_e value is the lowest out of all of the models, and indicates that this 1/y model has low variability. Additionally, the R^2 value of the 1/y model of 0.187 is slightly greater than that of the simple linear regression model R^2 of 0.1555, and only slightly lesser than the R^2 of the log(y) and log(x) & log(y) models, which indicates that more variability can be explained by a relationship with floor number. Additionally, the r value of 0.4324 is slightly better than that of the simple linear model which indicates a stronger correlation between floor number and 1/time. Finally, the slope of the 1/y model is statistically significant, as proven by our model utility test where our p-value of 2.08e-05 is less than our alpha of 0.05. The 1/y model is the best because the R^2 and r values are larger than the simple linear regression model and close to other transformed models. Additionally, the slope is significant, and the standard error is much smaller compared to any other models. Therefore we use this new regression: $1/\hat{y} = 0.4004 + 0.1056x$

Floor Number	Predicted Value
5	1.0771
6	0.9671
7	0.8775
8	0.8031
9	0.7403
10	0.6866
11	0.6402
12	0.5997

In this regression, when we extrapolate floor numbers, rather than going below 0, the predicted value always stays positive, as shown through our inputs of 5-12 floors into our model. Therefore it is confirmed that this model is more accurate.

Now that we have a model, we can use a 95% confidence interval on predicted values in the regression equation. We will be 95% confident that the true predicted value lies in the range of calculated values. We will use x values of 5-9.

<u>95% Confidence Interval for Predicted Values</u>

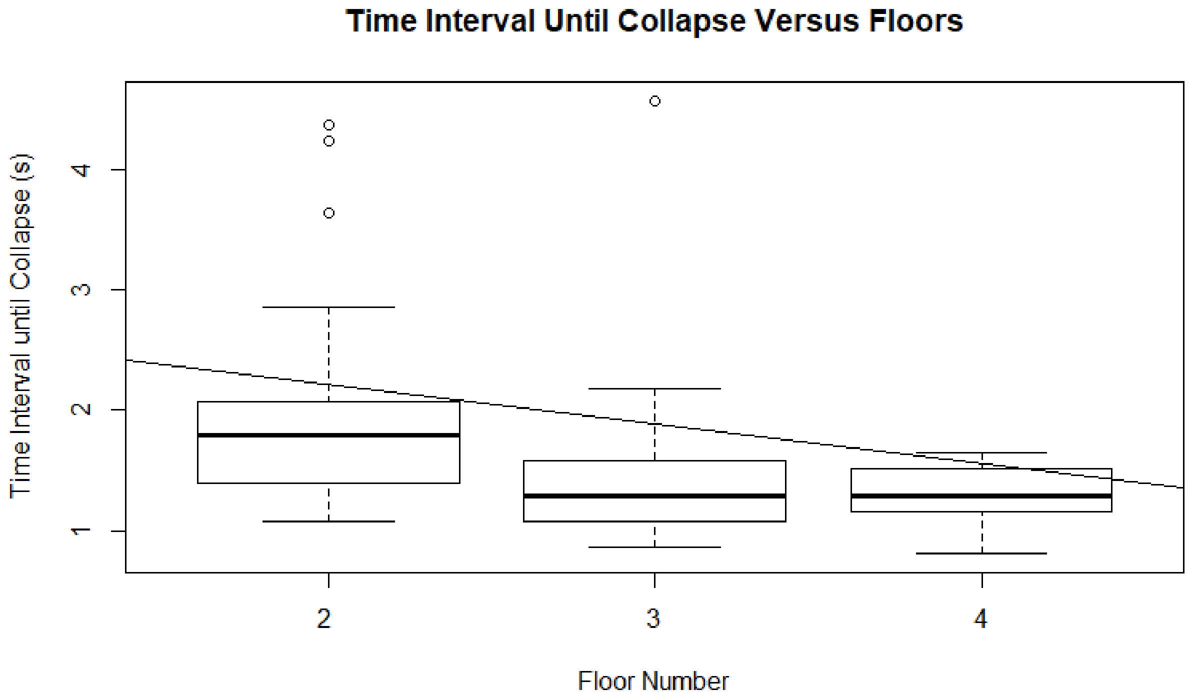
$\hat{y} \pm t^*(s_b)$ w/df = n-1

Sb=0.02347, df=89, t*=1.662

Floor Number	Predicted Value
5	1.0771±0.039 = (1.0381, 1.1161)
6	0.9671±0.039 = (0.9281, 1.0061)
7	0.8775±0.039 = (0.8385, 0.9165)
8	0.8031±0.039 = (0.7641, 0.8421)
9	0.7403±0.039 = (0.7013, 0.7793)

Graphs

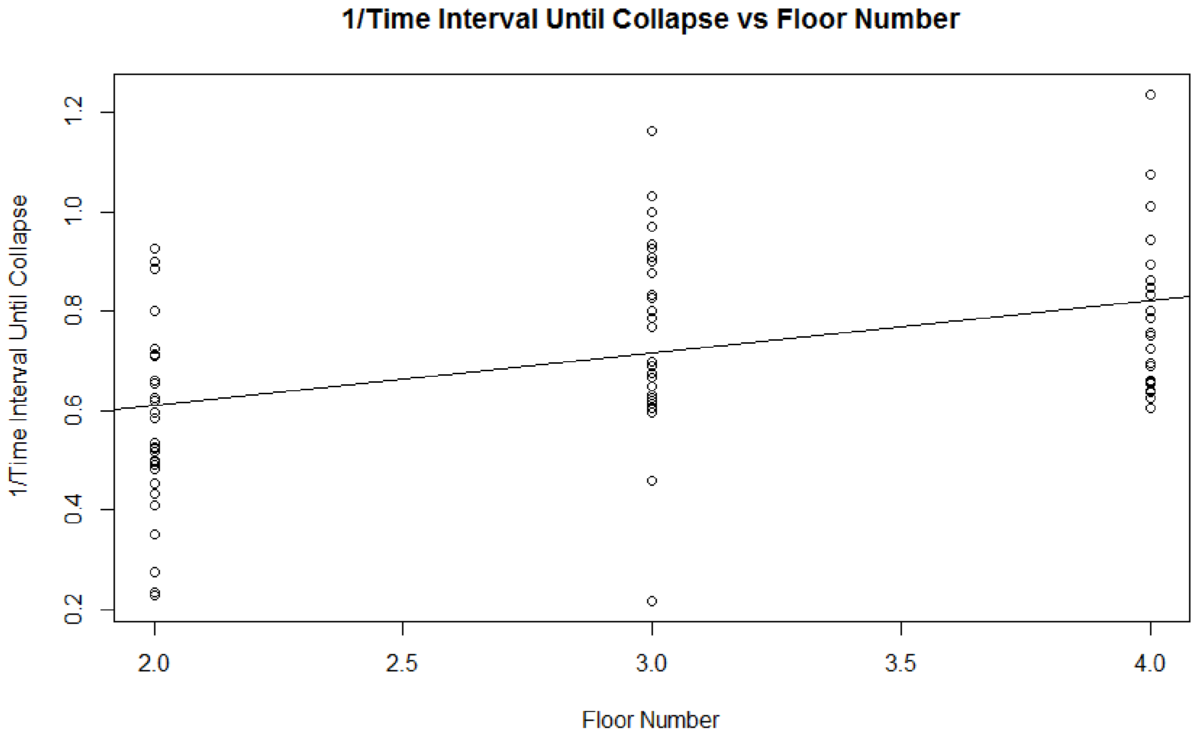
Simple Linear Regression Model



We made a boxplot regression graph in order to show the regression line of the data of the floor number vs time interval, as well as to display the spread of the data.

The regression line of $\hat{y}$=2.5455-0.3288x has a low R^2 value of 0.1555, meaning that only 15.55% of the variance within the time interval until building collapse can be explained by a linear relationship with floor number. The r-value of -0.3943 indicates a weak, negative, linear relationship between floor number and time interval and collapse. The large S_e value of 0.6327 indicates that there is a lot of variance in this model. These three factors show how this simple linear model was not the best model. Thus, we transformed the data.

Transformed Model Plot



This is the plot of the transformed model that we calculated, where we have a 1/y transformation and the equation: $1/\hat{y} = 0.4004 + 0.1056x$. The R^2 value of 0.187 indicates that 18.7% of variability within 1/time interval until collapse can be explained by a linear relationship with floor number. The r-value of 0.4324 of 0.4324 suggests a moderate positive linear relationship between floor number and 1/time interval until collapse. The low S_e value of 0.1818 indicates that there is not a lot of variance in this model. Additionally, the slope is statistically significant according to our model utility test, where our p value of 2.08e-05 is lower than the alpha value of 0.05. This is a better model than the simple linear regression one.

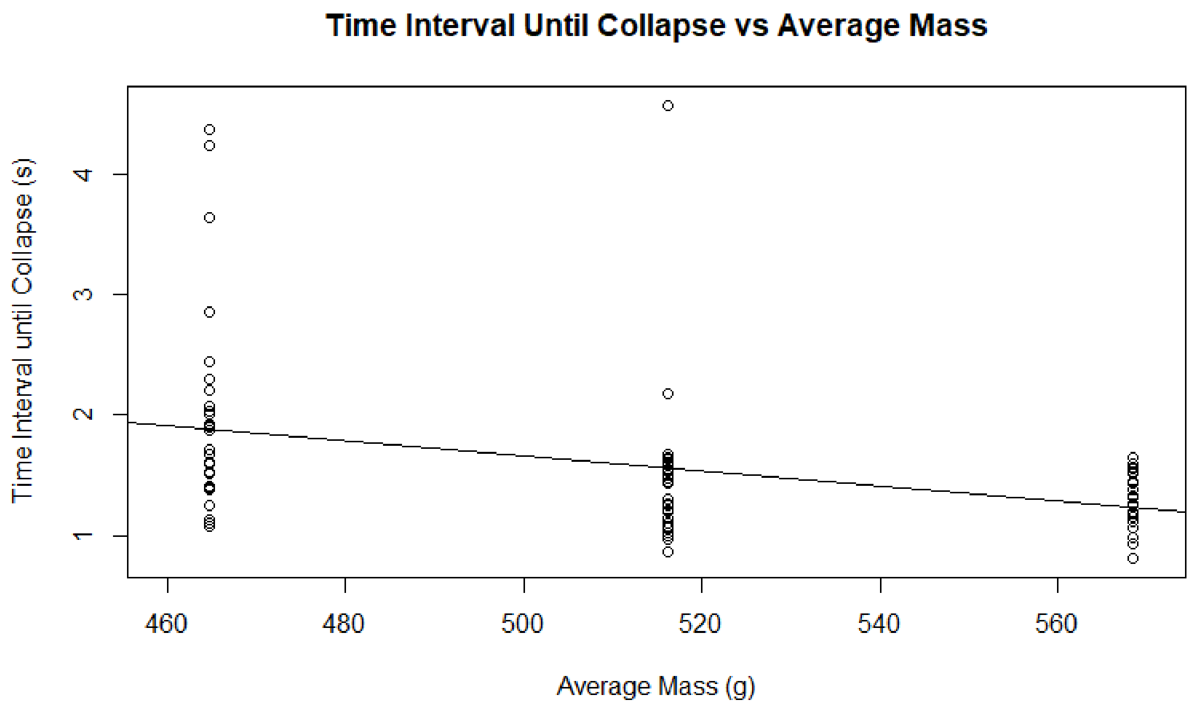
Analysis of the Spread

Floor	Min	Q1	Median	Q3	Max	Mean	Standard Deviation	IQR	Range
2	1.080	1.402	1.790	2.060	4.370	1.949	0.8444	0.6575	2.290
3	0.860	1.085	1.285	1.570	4.570	1.437	0.6572	0.485	3.710
4	0.810	1.165	1.295	1.495	1.650	1.291	0.2159	0.330	0.840

Each of our boxplots look relatively normally distributed, disregarding the outliers. Floor 2 has three outliers on the upper end, meaning that the boxplot would be skewed positively if these values were added. Floor 3 is slightly positively skewed as well, and has one upper end outlier. Finally, Floor 4 has no outliers and looks approximately normal. The means of the Floor 2-4 data decrease as floors increase; Floor 2 has a mean of 1.949, Floor 3 has a mean of 1.437 and Floor 4 has a mean of 1.291. This indicates that as the floor increases, the mean time interval until collapse decreases. Although the mean did decrease for all consecutive floors, the medians of the data were different. From floors 2-4, the median did not always decrease; Floor 2 had a median of 1.790, Floor 3 had a median of 1.285, and Floor 4 had a median of 1.295. Floor 4 actually had a higher median than that of Floor 3. This is because while Floor 3 had a positive skew with the mean of 1.437 being much higher than the median of 1.285, Floor 4 was normally distributed, with the median of 1.295 and mean of 1.291, and also because Floor 3 had larger spread than Floor 4 (larger standard deviation). In terms of standard deviation, once again, this number decreased as floors increased, showing a negative correlation; Floor 2 had a standard deviation of 0.8444, Floor 3 had a standard deviation of 0.6572, and floor 4 had a standard deviation of 0.2159. As the floor increased, the amount of variance decreased. The IQR supports this; it decreased from 0.6575 to 0.485 to 0.330 from floors 2-4. This shows how Floor 4 had the most compact data and had less spread. In terms of range, Floor 3 had the greatest with 3.710, followed by Floor 2 with 2.290, and Floor 4 with 0.840. The outliers in the data of Floors 2 and 3 caused an inflation of the range and caused it to be so large. Floor 4, with no outliers, had the smallest range.

Overall, Floor 4 had the least spread, with no outliers, a relatively normal distribution and the smallest IQR, range, and standard deviation. Floor 3 had the second least spread with 1 outlier, a positive skew, and the second smallest IQR and standard deviation. Finally, Floor 4 had the most spread with 3 outliers, a positive skew, and the largest IQR and standard deviation. The medians of the floor data decreased from 1.790 to 1.295 as the floor number increased from 2 to 4.

Average Mass Regression



Mass is directly related to floor number. We plotted time interval until collapse vs average mass to find an equation that would be able to predict the time interval until collapse of a tower that had different masses (this estimate is only for blocks in the 3x3 tower shape we mentioned before).

The regression equation was: $\hat{y} = 4.8303 + -0.0063x$

This is the same data as when we analyzed time interval until collapse vs floor number. Therefore the R^2 and r-value are still the same, with a low 0.1555 and 0.3943 respectively. Additionally, the S_e is still high, with a large 0.6328, and the slope is statistically significant with a p-value of 0.0001.

We compared the time interval to an average mass, because we only took three trials of the total mass after we were done testing; we did not record the mass of the tower every time we took one trial. Therefore, we must look at just the average.

This model is not the best to predict the time interval, and because this data is the same as our previously analyzed data, we know that the best transformation is a 1/y transformation. Thus our new model is: $1/\hat{y} = 0.4004 + 0.1055x$

Conclusion

We mostly accept our hypothesis. We thought that the time interval for building to collapse would decrease as the floor number increased; this was true. However, we also predicted that the differences would be statistically significant from each other, but this was not the case between floors 3 and 4. We were able to transform our data from a simple linear regression model that was not very accurate (low R^2, low r, high S_e), and found that a 1/y model worked the best to predict the time interval for collapse based on number of floors: $1/\hat{y} = 0.4004 + 0.1056x$. Additionally, we were able to generate estimates of extrapolated predicted values in this model by using a 95% confidence interval to generate intervals from which we were 95% confident that the true mean lied in. We also fit a model to predict time interval based on average mass, and once again, the 1/y model was the best. We received the regression equation: $\hat{y} = 4.8303 + -0.0063x$. One reason our results were not consistent with previous data was because we were using such a high frequency (over 4 Hz for each tower). This could have meant the frequency was too large for even taller buildings, which would normally be able to withstand short bursts of intense shaking ("Building"). Some potential confounding variables could be that we did not place each individual Jenga block in the same spot after it fell over, so this could have slightly altered the results if each Jenga block is not exactly the same. However, each block is very consistent with each other in size and dimension. Another potential confounding variable could have been the different sound sources that were playing during my tests. The sound could have slightly altered how quickly the tower fell down, which may have altered results. However, this would only be a miniscule difference. Another potential confounding variable could have been that because we were hand timing, the timing would have not been completely accurate compared to a computer. However, because we conducted 30 trials and randomized treatments, the effects of this are limited. Finally, our last potential confounding variable is that because we were turning the knob on the EME by hand, we might not have turned it to the exact same spot every trial, which would have resulted in different shake strengths and frequencies. Once again, we randomized treatments and had 30 trials, so the effects of this are minimized. In the future, we would like to see how the foundation and the depth of the foundation affect the integrity of the building. In this experiment, we simply placed the building on our plastic holding container. However, we could experiment with different materials to place the tower in/on in order to maximize the time interval it takes for it to fall down. Additionally, we would like to see how the various floors react to different shake strengths (therefore different frequencies of shaking) to see how the time interval changes.

Appendix

Raw Data

Time Interval: Floor 2 (s)	Time Interval: Floor 3 (s)	Time Interval: Floor 4 (s)
1.38	1.65	1.56
1.93	1.58	1.2
1.51	2.18	1.32
1.9	1.63	1.6
4.23	1.45	0.81
1.61	1.3	1.52
2.44	1.21	1.32
2.01	4.57	1.33
2.85	1.61	0.99
2	1.68	1.57
1.6	1.45	1.25
1.68	1.08	1.44
1.08	0.97	1.65
1.11	1	1.27
1.71	1.11	0.93
2.07	1.08	1.18
1.53	1.25	1.06
1.91	1.5	1.45
1.13	1.07	1.38
2.2	1.48	1.06
2.3	1.6	1.27
3.63	1.03	1.18
1.25	1.2	1.12
1.41	0.86	1.53
1.25	1.27	1.51
1.87	1.54	1.06
1.08	1.43	1.52
1.4	1.1	1.32
4.37	1.08	1.16
2.03	1.14	1.18

Frequency: Floor 2 (hz.)	Frequency: Floor 3 (hz.)	Frequency: Floor 4 (hz.)
4.4	4.2	4
4.5	4.1	3.9
4.3	4.4	3.9

Mass: Floor 2 (g)	Mass: Floor 3 (g)	Mass: Floor 4 (g)
463.4	515.02	567.11
466.21	517.84	569.94
464.33	516	568.11

Works Cited

"Buildings and Earthquakes—Which Stands? Which Falls?" IRIS and the University of Portland, n.d. Web. 7 May 2017.

Kishida, Hideaki. "Damage to reinforced concrete buildings in Niigata city with special reference to foundation engineering." Soils and Foundations. The Japanese Geotechnical Society, 29 Feb. 2008. Web. 8 May 2017.