Name: Charlotte Weinstein

UmichID: cbweinst

What dataset are you working with: san\_andreas

List 3 questions that you can ask with your dataset.

Q1: Are people who have experienced an earthquake more likely to be worried about earthquakes?

Q2: Do people who are worried about the “Big One” think that it will occur in their lifetime?

Q3: Are people familiar with the Yellowstone Supervolcano more or less likely to be worried about the “Big One”?

List the associated null hypothesis for each question:

Q1: People who have experienced an earthquake are not more worried about earthquakes than people who haven’t experienced an earthquake.

Q2: People who are worried about the “Big One” do not think that it will occur in their lifetime.  
Q3: People familiar with the Yellowstone Supervolcano are less worried about the “Big One” than people who aren’t familiar with the Yellowstone Supervolcano.

What statistical test(s) will you use to answer each of the questions:

Q1: chi-square

Q2: chi-square

Q3: chi-square

Make a visual plot showing the relationship that you will analyze statistically (e.g. boxplot for t-test or ANOVA; scatterplot for regression; table for chi-square).

Q1:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Experience? | Not at all worried | Not so worried | Somewhat worried | Very worried | Extremely worried |
| No | 175 | 106 | 52 | 9 | 18 |
| Yes, one or more minor ones | 149 | 172 | 113 | 28 | 21 |
| Yes, one or more major ones | 30 | 49 | 55 | 16 | 13 |

Q2:

|  |  |  |
| --- | --- | --- |
|  | False | True |
| Not at all worried | 254 | 80 |
| Not so worried | 194 | 116 |
| Somewhat worried | 87 | 139 |
| Very worried | 19 | 56 |
| Extremely worried | 23 | 45 |

Q3:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Not at all worried | Not so worried | Somewhat worried | Very worried | Extremely worried |
| Not at all familiar | 108 | 74 | 52 | 14 | 22 |
| Not so familiar | 57 | 92 | 39 | 14 | 17 |
| Somewhat familiar | 96 | 80 | 74 | 19 | 11 |
| Very familiar | 38 | 37 | 41 | 20 | 3 |
| Extremely familiar | 32 | 26 | 15 | 7 | 13 |

Do your data meet the assumptions required for the statistical test you want to run? Please state the assumptions you examined and whether or not your data meet those assumptions:

Q1: Assumption 1: Each observation is independent of all the others – this is a general assumption, based on proper survey technique.

Q2: Same as assumption for Q1.

Q3: Same as for Q1 & Q2

Run the statistical test! Put your results here:

Q1: Pearson's Chi-squared test

data: san\_andreas$experience and san\_andreas$worry\_general

X-squared = 71.957, df = 8, p-value = 2.001e-12

Q2: Pearson's Chi-squared test

data: san\_andreas$worry\_bigone and san\_andreas$will\_occur

X-squared = 130.51, df = 4, p-value < 2.2e-16

Q3: Pearson's Chi-squared test

data: san\_andreas$fam\_yellowstone and san\_andreas$worry\_bigone

X-squared = 58.018, df = 16, p-value = 1.125e-06

Interpret your results!

Q1: Our p-value is significant, so we can reject the null hypothesis that people who have experienced an earthquake are not more worried about earthquakes than people who haven’t experienced an earthquake.

Q2: Like in Q1, our p-value is significant, so we can reject the null hypothesis that people who are worried about the “Big One” do not think that it will occur in their lifetime.

Q3: Similar to Q1 and Q2, our p-value is very small, so our results are significant and we can reject the null hypothesis that people familiar with the Yellowstone Supervolcano are less worried about the “Big One” than people who aren’t familiar with the Yellowstone Supervolcano.