1. Problem 1

A federal agency is deciding which of two waste dump projects to investigate. A top administrator estimates that the probability of federal law violations is 0.30 at the first project and 0.25 at the second project. Also, he believes the occurrences of violations in these two projects are disjoint.

a. What is the probability of federal law violations in the first project or in the second project?

b. Given that there is not a federal law violation in the first project, find the probability that there is a federal law violation in the second project.

c. In reality, the administrator confused disjoint and independent, and the events are actually independent. Answer parts a and b with this correct information.

2. Sample R Problem: Tornadoes in Minnesota

The tornadoes dataset (available at `http://stat.umn.edu/~wuxxx725/data/tornadoes.csv`) is consisted of several measurements for the 1,371 tornadoes that had made touchdowns in Minnesota between 1950 and 2006 inclusive. (Source: NOAA/NWS Storm Prediction Center)

In this problem, we consider only the following variables:

- Sig: A logical variable indicating whether the tornado is a significant tornado, that is, whether its Fujita scale is 2 or above.
- Area: A categorical variable indicating the region of Minnesota where the touchdown occurred (NW = Northwest, N = North Central, NE = Northeast, W = West Central, C = Central, E = East Central, SW = Southwest, S = South Central, SE = Southeast. Source: NWS Climate Prediction Center).

We investigate the relationship between Area and Sig to determine whether the geographical areas and the strengths of tornadoes are associated.

a. Generate a cross tabulation of Sig by Area. Note that the table command does not give you the row, column, or overall total. You need to either add up each row and column by hand or use the margin.table() command to obtain the sums.
(Note: There were 8 tornadoes for which the Sig variable is missing. Exclude them for this and subsequent parts of this problem.)

b. For a randomly selected tornado that had occurred in Minnesota, what is the sample space of the possible outcomes for Sig? What about for Area?

c. Let E and T denote the events that the tornado made the touchdown in East Central Minnesota and that the tornado is a significant tornado, respectively. Using the data above to estimate the probabilities Pr(E) and Pr(T).

d. Estimate the probability that a Minnesota tornado made the touchdown in East Central Minnesota and was significant.

e. Estimate the probability that a Minnesota tornado made the touchdown in East Central Minnesota or was significant (or both).

f. Based on your answers from parts c. and d., are events E and T independent? Please use statistical reasoning and a formula to justify your answer.

g. Based on your answers from parts c. and d., are events E and T disjoint? Justify your response.

h. Estimate the probability that a Minnesota tornado was significant, given that it had made the touch down in East Central Minnesota. Comment on whether East Central Minnesota has a higher proportion of tornadoes that are significant compared to the whole state of Minnesota.