

Highways Classwork 1: Contingency tables, Independence

Use the Highways article to answer the following questions.

- 11,337,933 people make up 3.7% of the total US population. Using that information, what is the total population of the US?
- Consider the two populations labeled Non-Hispanic White and Non-Hispanic Black. Of these two, which population had more people living within 150 meters of a major highway? Which population had a larger percentage of people living within 150 meters of a major highway? Which of those statistics do you think is more meaningful in this conversation? Why?
- Approximately how many people in the US speak primarily Spanish in their households?
- How do the authors define the "Poor" characteristic? What how many people in the US, in total, are classified as "Poor" under this definition?
- Using the information from the table found in the article, fill in the blanks to complete the following contingency table, **Table 1**. This compares people's proximity to major highways to the language predominantly spoken in their home. Each percentage shown is of the row total.

Language	Proximity		Total
	< 150 m	> 150 m	
English only	7,513,304 (3.3%)	220,162,575 (96.7%)	227,675,879 (100%)
Spanish	1,805,261 (5.1%)		
Other	1,059,572 ((4.9%)		
Total	10,378,137		

- Notice that the number you found for "total" in the first column (10,378,137) differs a bit from the "Total" number found in the table in the article (11,337,933). Why do you think this is?
- Repeat each of the counts from Table 1. Then divide each by the **table total** and express as a percent. Record below in **Table 2**.

Language	Proximity		Total
	< 150 m	> 150 m	
English only	7,513,304 (____ %)	220,162,575 (____ %)	227,675,879 (____ %)
Spanish	1,805,261 (____ %)		
Other	1,059,572 ((____ %)		
Total	10,378,137 ((____ %)		

8. What proportion of the US population primarily speaks Spanish in their home?

9. What proportion of predominantly Spanish-speakers live close to a major highway?

10. Of the people who live close to a major highway, what proportion of them are predominantly Spanish-speakers?

11. Two variables are **independent** if the conditional distribution for one variable remains unchanged with respect to the other variable. We will calculate what the distribution *would be* if we were to **assume independence**. To do this, first enter the marginal distributions (Total row at the bottom, Total column on the right), and the table total from Table 2. Then ignore the other values in Table 2; instead, fill in the value for (Row Total) X (Column Total)/(Table Total) . These are *expected counts* under the assumption of independence. Round values to 1 decimal place. Record below in **Table 3**.

Language	Proximity		Total
	< 150 m	> 150 m	
English only			227,675,879
Spanish			
Other			
Total	10,378,137		

12. Compare the counts in Table 2 and the expected counts in Table 3. Are they reasonably close? Do you think there is some evidence that Language and Proximity are independent? Explain briefly.