MEMORANDUM

To:	Carole Voulgaris, Course Instructor, SES 5215
From:	Ramona Quimby, Student, SES 5215
Date:	October 28, 2022
Subject:	Assignment 2, Testing relationships

The purpose of this memo is to test for relationships among variables relevant to the question:

What is the effect of a Boston census tract's median age on the number of crashes per square kilometer that occur there, after accounting for the effects of poverty and population density?

I will be addressing this question using crash data from Vision Zero Boston (https://data.boston.gov/dataset/vision-zero-crash-records) and demographic data from the 2016-2020 American Community Survey. My dataset includes 187 census tracts in Boston (excluding tracts with zero households), shown in Figure 1 below.

Figure 1: Census tracts in the study area

I am testing the relationship between crash density (the number of crashes recorded in each census tract between January 1, 2015 and September 30, 2022, divided by the area (square kilometers) of land within the census tract) and each of the following variables.

1. Median age: The median age of census tract residents.
2. Population density: The number of residents per square kilometer of land area in each census tract.
3. Majority poverty: A categorical variable indicating whether the majority of census tract residents have income below the poverty level.

Relationship between median age and crash density

The relationship between a census tract's median age and its crash density is illustrated in Figure 2. The y-axis in Figure 2 is log-transformed. The 95-percent confidence interval for the correlation between a tract's median age and its crash density is between -0.24 and 0.04. The 95-percent confidence interval for the correlation between median age and the log of a tract's crash density is between -0.27 and 0.01. Both intervals include zero, so we cannot be 95-percent confident that there is a linear or log-linear relationship between median age and crash density.

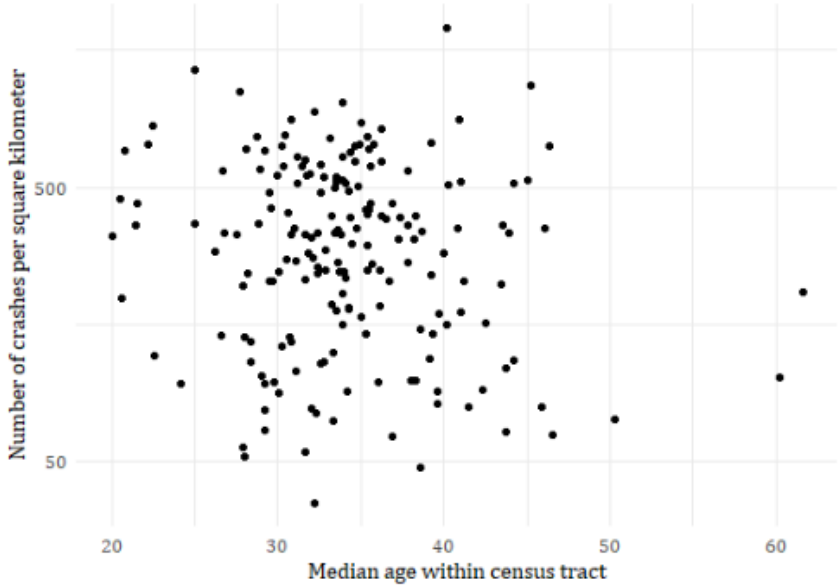

Figure 2: Relationship between median age and crash density

Relationship between population density and crash density

The relationship between a census tract's population density and its crash density is illustrated in Figure 3 (both axes are log-transformed). The 95-percent confidence interval for the correlation between the log of tract's population density the log of its crash density is between 0.16 and 0.43. This suggests that we can be 95-percent confident that there is a positive relationship between crash density and population density.

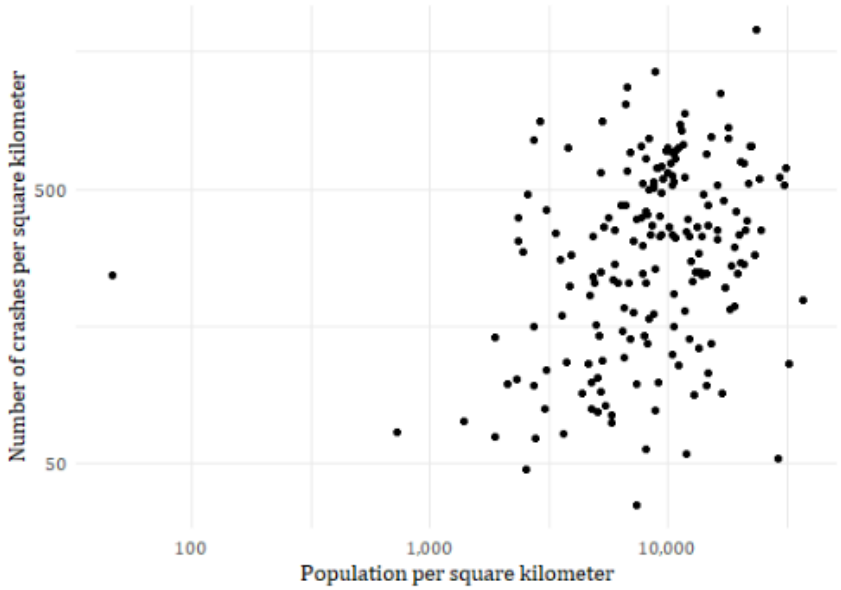

Figure 3: Relationship between population density and crash density

Relationship between poverty and crash density

Figure 3 shows the average crash densities within tracts where the majority of household incomes are below the poverty level and within tracts where the majority of household incomes are above the poverty level. Error bars represent 95-percent confidence intervals. The 95-percent confidence interval for the average crash density in tracts with a majority of households earning above the poverty level is between 312 and 391 crashes per square kilometer. The 95-percent confidence interval for the average crash density in tracts with a majority of households earning below the poverty level is 349 to 686 crashes per square kilometer. The difference in means between these two groups of tracts has a p-value of 0.058, which indicates the difference is not significant a 95-percent confidence level (although it would be significant at a 90-percent confidence level).

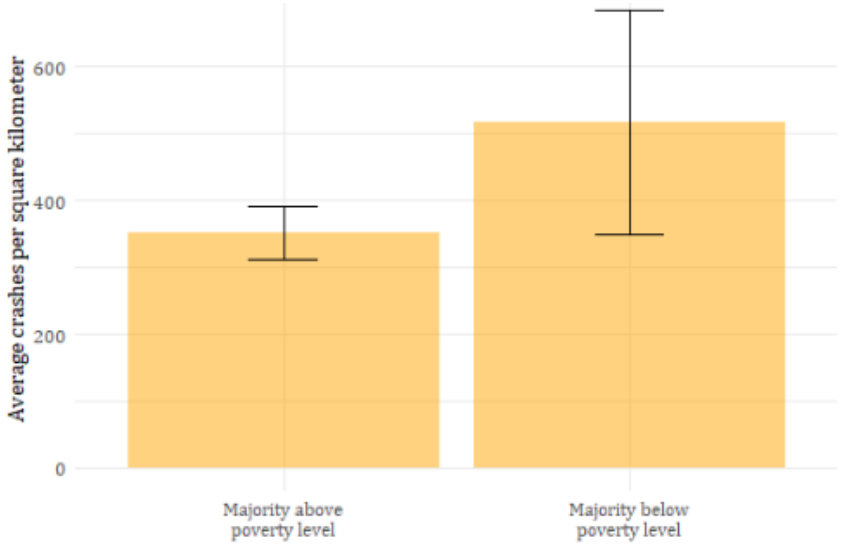

Figure 3: Average crash density by tract poverty status