

## 2017 Appraisal Values

The file [appraisal\\_2017.csv](#) contains the 2017 appraised value (total\_appr) and square footage (finished\_area) for a random sample of 1000 houses in Davidson County.

- 1) Create a scatterplot of total\_appr vs finished\_area.
- 2) By inspecting the scatterplot, describe the relationship between total\_appr and finished\_area. Is the direction of association positive or negative? Is the relationship linear? How strong is the relationship?
- 3) Do you see any points which might be considered outliers? Investigate those points.
- 4) Find the correlation between total\_appr and finished\_area. How strong is the relationship between the two variables?

## Crashes

The file [crashes\\_subset.csv](#) contains all reported crashes in Davidson County in 2018 which were classified as either head-on, front-to-rear, or a sideswipe where the cars were moving in the same direction.

- 1) Find the count of crashes by collision type. Create a bar plot showing the this count.
- 2) Create a boxplot showing number of injuries vs collision type description. What do you notice?
- 3) Find the average number of injuries per crash by collision type. Which type of crash has the highest average number of injuries. Create a box plot to show the distribution of injuries by collision type.
- 4) Compare the rate of hit and run crashes by category. Create a bar plot to show what you find. What do you notice?