

1. `names(houses)`
2. houses sold in this New York county
3. `nrow(houses)`
4. `gf_histogram(~ price, data = houses)`
5. unimodal, right-skewed
6. The mean is larger, due to the right skew.
7. `mean(~ price, data = houses)`
8. Our data is a sample from this larger population. So, the mean of Problem 7 is a sample statistic, denoted as  $\bar{x}$  (not  $\mu$ ).
9. `qdata(~ price, data = houses)`
10. B
11. It's about  $265 - 145 \sim \$120K$
12. mean, range, std. deviation
13. `filter(houses, price < 600000)`
14. Possibilities include:

Is there a difference in mean price for waterfront houses than for others?

Is there an association between price and the number of bedrooms?

↑  
or heating system  
or living area