* Presentation: What is a hypothesis and how do you test it?
  + Introduce the concept of hypothesis testing in statistical analysis and how to craft one
    - Hypothetical about how to measure when things are on average different…
      * Demo: How would we gather this data in Excel?
        + Go into the Housing dataset. We will see if homes with air conditioning have a higher sale price.
        + Use PivotTables so it’s easy

Make sure that totals are checked-off

* + - * + Now I am going to copy and paste one more PivotTable, this time I will use it to graph the distributions.
        + Drill: do the same for the full, finished basement.
* We want to rigorously test this now.
* It has to be falsifiable
* And it’s up to us to prove it.
* So, let’s translate this to math.
* Okay, we are ready to
  + - How can we test this for certain when we don’t have all the data?
  + Thanks to the central limit theorem, we can speak probabilistically about a sample’s mean. Here’s how it works.
    - We will calculate a test statistic and compare it to our critical value.
    - So let’s go into some math to do this.
* We can calculate our sample mean
  + We then want to compare it to the population mean
  + What we’ll do is compare some test statistic and see, how generalizable is this result?
    - How likely is our value to fall within our bounds of confidence?
    - Based on this difference we will get a p-value. This tells us statistical significance.
  + But we also want to establish substantive significance.
    - What are those bounds, exactly?
    - The confidence interval will tell us that.
    - With 95% confidence?

So, the big issue with the central limit theorem is sometimes we don’t know when a sample is “big enough” so if something is really lopsided then we may need pretty big samples.

If the samples we are working with are close to normal then we can say that our results will be more robust. If you appear to have something of a bell curve then you should be fine.