

First questions to ask yourself:

1. What are the observational units (people or things we took measurements on)?
2. What are the variables, and is each categorical or quantitative?
3. Are there one or two populations? (are we comparing proportions or means among two populations?)
4. If we are comparing means in two populations, are the data paired or unpaired?

	Assumptions to Check
Categorical (one)  Population proportion	<p>Independence</p> <ul style="list-style-type: none"> <li>- Randomization</li> <li>- Sample size (less than 10 percent of population size)</li> </ul> <p>Two possible outcomes for each trial.</p> <p>Same probability of success for each trial</p> <p>Sample size big enough: <math>np \geq 10</math> and <math>n(1-p) \geq 10</math></p>
Categorical (two populations)  Difference between population proportions	<p>All of the same conditions as above for each group</p> <p>+</p> <p>Independence between two groups</p>
Quantitative (one)  Mean for Quantitative	<p>Independence</p> <ul style="list-style-type: none"> <li>- Randomization</li> <li>- Sample size (less than 10 percent of population size)</li> </ul> <p>Distribution of values to be nearly normal (need the mean to be a good summary of center)</p> <ul style="list-style-type: none"> <li>- Unimodal</li> <li>- Symmetric</li> </ul> <p>Sample Size big enough (aiming for the sampling distribution of the sample mean to be nearly normal – how big depends on how far from normal)</p>
Quantitative (two populations, unpaired)  Difference between population means	<p>All of the same conditions as above for each group</p> <p>+</p> <p>Independence between two groups</p>
Quantitative (two populations, paired – same variable measured multiple times for each person or subject)  Difference between population means	<ol style="list-style-type: none"> <li>1. Differences between paired observations are independent across different pairs.</li> <li>2. Differences between paired observations follow a nearly normal distribution</li> <li>3. Sample Size big enough</li> </ol>
Quantitative (one population, two quantitative variables measured)  Regression/linear association	<p>O'LINE</p> <p>no Outliers</p> <p>Linear relationship (straight enough?)</p> <p>Independent observations</p> <p>Normally distributed residuals</p> <p>Equal variance of residuals (does the plot thicken?)</p>