Scenario	One Categorical Response	Two Categorical Variables	One Quantitative Response OR Paired Differences	Two Quantitative Variables	Quant. Response and Categ. Explanatory (independent samples)
Type of plot	Bar plot	Segmented bar plot, Mosaic plot	Dotplot, histogram, boxplot	Scatterplot	Side-by-sided boxplots, Stacked dotplots or histograms
Summary measure	Proportion	Difference in proportions	Mean or Mean Difference	Slope or correlation	Difference in means
Parameter notation	π	$\pi_1 - \pi_2$	μ or μ_d	β_1 or ρ	$\mu_1 - \mu_2$
Statistic notation	$\hat{p}$	$\widehat{p_1} - \widehat{p_2}$	$\bar{x}$ or $\bar{x}_d$	b_1 or r	$\bar{x}_1 - \bar{x}_2$
Null hypothesis	$H_0: \pi = \pi_0$	$H_0: \pi_1 - \pi_2 = 0$	$H_0: \mu = \mu_0$ or $H_0: \mu_d = 0$	$H_0: \beta_1 = 0$ or $H_0: \rho = 0$	$H_0: \mu_1 - \mu_2 = 0$
Conditions for simulation methods	Independent cases;	Independence (within and between groups);	Independent cases;	Independent case, Linear form;	Independence (within and between groups);
Simulation test (how to generate a null distn) p-value = proportion of null simulations at or beyond (H_A direction) the observed statistic	Spin spinner with probability equal to π_0, n times or draw with replacement n times from a deck of cards created to reflect π_0 as probability of success. Plot the proportion of successes. Repeat 1000's of times. Centered at π_0	Label cards with response values from original data; mix cards together; shuffle into two new groups of sizes n_1 and n_2. Plot difference in proportion of successes. Repeat 1000's of times. Centered at 0.	Shift the original data by adding $(\mu_o - \bar{x})$ or $(0 - \bar{x}_d)$. Sample with replacement from the shifted data n times. Plot sample mean. Repeat 1000's of times. Centered at μ_0 (single mean) or 0 (paired mean difference).	Hold the x values constant; shuffle y's to new x's. Find the regression line for shuffled data; plot the slope or the correlation for the shuffled data. Repeat 1000's of times. Centered at 0.	Label cards with response variable values from original data; mix cards together; shuffle into two new groups of sizes n_1 and n_2. Plot difference in means. Repeat 1000's of times. Centered at 0.
Bootstrap CI (how to generate a boot. distn) X% CI: $(\frac{1-X}{2}\%tile, \left(X + \frac{1-X}{2}\right)\%tile)$	Label n cards with the original responses. Randomly draw with replacement n times. Plot the resampled proportion of successes. Repeat 1000's of times. Centered at $\hat{p}$.	Label $n_1 + n_2$ cards with the original responses. Randomly draw with replacement n_1 times from group 1 and n_2 times from group 2. Plot the resampled difference in proportion of successes. Repeat 1000's of times. Centered at $\widehat{p_1} - \widehat{p_2}$.	Label n cards with the original responses. Randomly draw with replacement n times. Plot the resampled mean. Repeat 1000's of times. Centered at $\bar{x}$ or $\bar{x}_d$.	Label n cards with the original (explanatory, response) values. Randomly draw with replacement n times. Plot the resampled slope or correlation. Repeat 1000's of times. Centered at b_1 or r.	Label $n_1 + n_2$ cards with the original responses. Randomly draw with replacement n_1 times from group 1 and n_2 times from group 2. Plot the resampled difference in means. Repeat 1000's of times. Centered at $\bar{x}_1 - \bar{x}_2$.
Theory-based distribution	Standard Normal	Standard Normal	t- distribution with $n - 1$ df	t- distribution with $n - 2$ df	t- distribution with min of n_1-1 or n_2-1 df
Conditions for theory-based hypothesis tests	Independent cases; Number of expected successes and number of expected failures both at least 10. $\pi_o * n \geq 10, (1 - \pi_o) * n$	Independence (within and between groups); Number of expected successes and number of expected failures in each group is at least 10. $\hat{p}_{pool} * (n_1) \geq 10, (1 - \hat{p}_{pool}) * (n_1)$ $\hat{p}_{pool} * (n_2) \geq 10, (1 - \hat{p}_{pool}) * (n_2)$	Independent cases; $n < 30$ with no clear outliers OR $30 \leq n < 100$ with no extreme outliers OR $n \geq 100$	Independent cases; Linear form; Nearly normal residuals; Variability around the regression line is roughly constant.	Independent cases (within and between groups); In each sample, $n < 30$ with no clear outliers OR $30 \leq n < 100$ with no extreme outliers OR $n \geq 100$
Theory-based standardized statistic (test statistic)	$z = \frac{\hat{p} - \pi_0}{SE_0(\hat{p})}$ $SE_0(\hat{p}) = \sqrt{\frac{\pi_0 \times (1 - \pi_0)}{n}}$	$z = \frac{\widehat{p_1} - \widehat{p_2}}{SE_0(\widehat{p_1} - \widehat{p_2})}$ $SE_0(\widehat{p_1} - \widehat{p_2}) = \sqrt{\widehat{p_{pool}} \times (1 - \widehat{p_{pool}}) \times \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$, where $\widehat{p_{pool}} = \frac{total\ successes}{total\ sample\ size} = \frac{n_1 \times \widehat{p_1} + n_2 \times \widehat{p_2}}{n_1 + n_2}$	$t = \frac{\bar{x} - \mu_0}{SE(\bar{x})}$ $SE(\bar{x}) = \frac{s}{\sqrt{n}}$	$t = \frac{b_1}{SE(b_1)}$ $SE(b_1)$ is the reported standard error (std. error) of the slope term in the lm() output from R.	$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{SE(\bar{x}_1 - \bar{x}_2)}$ $SE(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
Conditions for theory-based confidence intervals	Independent cases; Number of successes and number of failures in the sample both at least 10.	Independence (within and between groups); Number of successes and number of failures in EACH sample all at least 10. (All four cell counts at least 10.)	Independent cases; $n < 30$ with no clear outliers OR $30 \leq n < 100$ with no extreme outliers OR $n \geq 100$	Independent cases; Linear form; Nearly normal residuals; Variability around the regression line is roughly constant.	Independent cases (within and between groups); In each sample, $n < 30$ with no clear outliers OR $30 \leq n < 100$ with no extreme outliers OR $n \geq 100$
Theory-based confidence interval	$\hat{p} \pm z^{\star} \times SE(\hat{p})$ $SE(\hat{p}) = \sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}}$	$\widehat{p_1} - \widehat{p_2} \pm z^{\star} \times SE(\widehat{p_1} - \widehat{p_2})$ $SE(\widehat{p_1} - \widehat{p_2}) = \sqrt{\frac{\widehat{p_1} \times (1 - \widehat{p_1})}{n_1} + \frac{\widehat{p_2} \times (1 - \widehat{p_2})}{n_2}}$	$\bar{x} \pm t^{*} \times SE(\bar{x})$ $SE(\bar{x}) = \frac{s}{\sqrt{n}}$	$b_1 \pm t^{*} \times SE(b_1)$ $SE(b_1)$ is the reported standard error (std. error) of the slope term in the lm() output from R.	$\bar{x}_1 - \bar{x}_2 \pm t^{*} \times SE(\bar{x}_1 - \bar{x}_2)$ $SE(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$