Data to be studied or compared	Types (and number) of Variables	Example of Hypotheses (Null and Alternative)	Test Statistic (Z-score, T-score etc and formula)	Large Sample			Small Sample		
				Type of Test	**Associated Distribution and parameters**	**Necessary Conditions**	**Type of Test**	**Associated Distribution and parameters**	**Necessary Conditions**
Mean Value	1 numerical variable	$H_0: \mu = 100$ $H_A: \mu \neq 100$	$Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$	Z-test for a mean	Normal Distribution	▪ n>30 ▪ Independent Observations ▪ No strong Skew	T-test for a mean	T-distribution parameter: degrees of freedom (n-1)	▪ Independent Observations ▪ Nearly Normal distribution
Difference of Means (paired data)	2 numerical variables (from each observation)	$H_0: \mu_{diff} = 0$ $H_A: \mu_{diff} \neq 0$	$Z = \frac{\overline{x_{diff}} - 0}{\sigma_{diff} / \sqrt{n}}$	Z -test for a paired difference of means	Normal Distribution	▪ n>30 ▪ Independent Observations ▪ No strong Skew	T-test for a paired difference of means	T-distribution parameter: degrees of freedom (n-1)	▪ Independent Observations ▪ Nearly Normal distribution
Difference of Means (unpaired data)	1 numerical variable, one 2 level categorical variable	$H_0: \mu_1 - \mu_2 = 0$ $H_A: \mu_1 - \mu_2 \neq 0$	$Z = \frac{\overline{x_1} - \overline{x_2} - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	Z -test for difference of means	Normal Distribution	▪n_1,n_2>30 ▪ Independent Observations ▪ No strong Skew	T-test for difference of means	T-distribution parameter: degrees of freedom min(n_1-1,n_2-1)	▪ Independent Observations ▪ Nearly Normal distribution
Comparison of Multiple Means	1 numerical variable, one categorical variable with multiple levels	H_0: All means equal H_A: At least one mean different from others	$F = \frac{MSG}{MSE}$ (Not on exam)	ANOVA	F-distribution parameters: degrees of Freedom Group and degrees of freedom error	▪ Independent Observations ▪ Nearly Normal distributions ▪ Constant Variance			
Proportion (counts) of data in one of two categories.	1 categorical variable (2 levels)	$H_0: \hat{p} = 0.4$ $H_A: \hat{p} \neq 0.4$	$Z = \frac{p_O - p_E}{\sqrt{\frac{p_E(1-p_E)}{n}}}$ p_E = Expected Proportion	Z-test for a proportion	Normal Distribution	▪ Independent Observations ▪ Expected counts at least 10	Simulation	Simulated Distribution	▪ Independent Observations

Difference of two proportions	2 categorical variables (both 2 level)	$H_0: \hat{p}=0.4$ $H_A: \hat{p}_1-\hat{p}_2\neq 0$	$Z=\frac{(p_1-p_2)-0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1}+\frac{\hat{p}(1-\hat{p})}{n_2}}}$ ($\hat{p}$ = pooled proportion, use when testing equality of proportions, otherwise use p_1, p_2 respectively) Z-Test for a difference of proportions	Normal Distribution	▪Independent observations ▪ Expected counts at least 10	Simulation	Simulated Distribution	▪ Independent Observations
Counts of data in more than two categories	1 categorical variable (multiple levels)	H_0: Counts match expected distribution H_A: Counts differ from expected distribution	$\chi=\sum\frac{(obs_i-\exp_i)^2}{\exp_i}$ Observed and expected counts from distribution. Chi Squared test for goodness of fit.	Chi Squared Distribution Parameter: Degrees of Freedom (# of bins-#of constraints -1)	▪Independent observations ▪ Expected counts at least 5 ▪ At least 2 degrees of Freedom	Simulation	Simulated Distribution	▪ Independent Observations
Relationship between counts of two different variables	2 categorical variables, at least one of which is multiple levels	H_0: Categorical variables are independent of each other H_A: Categorical variables are dependent, counts vary between rows/columns	$\chi=\sum\frac{(obs_i-\exp_i)^2}{\exp_i}$ Where the expected value of an entry is $\frac{(row\,total)\times(column\,total)}{table\,total}$ Chi squared test for independence	Chi Squared Distribution Parameter: Degrees of Freedom: (#rows -1) *(#cols-1)	▪Independent observations ▪ Expected counts at least 5 ▪ At least 2 degrees of Freedom	Simulation	Simulated Distribution	▪ Independent Observations
Correlation between two numerical values	2 Numerical Variables	$H_0: \beta_1=0$ $H_0: \beta_1\neq 0$	$\beta_1=R\frac{s_y}{s_x}$ $R=\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}}$ $\beta_0=\bar{y}-\beta_1\bar{x}$ T-score from computer output	T-test for linear regression	▪Linear Data ▪Nearly Normal Residuals ▪ Constant Variablilty			