Variable(s) of Interest	Parameter of Interest	Statistic of Interest	Descriptive Method(s)	Inferential Method(s)	Assumption(s) for Inferential Methods
Single Categorical Variable **(Binary – 2 categories)**	True Population Proportion (π)	Sample Proportion ($\hat{p}$)	▪ Report $\hat{p}$ ▪ Bar chart	▪ Simulation ▪ Binomial test ▪ CI for π	Check the four conditions for the binomial
Single Categorical Variable **(More than two categories)**	True Population Proportion $(\pi_1, \pi_2, ...)$	Sample proportions $(\hat{p}_1, \hat{p}_2 \ldots)$	▪ Report $\hat{p}_1, \hat{p}_2$ ... ▪ Stacked bar chart	▪ Chi-square Goodness of Fit	▪ Observations are independent ▪ EXPECTED counts should be greater than 5
Two Categorical Variables **(in general)**	True Conditional Population Proportions $(\pi_{1\|G1}, \pi_{1\|G2}, ...)$	Sample Proportions $(\hat{p}_{1\|G1}, \hat{p}_{1\|G2} \ldots)$	▪ Report sample proportions ▪ Contingency table ▪ Stacked/Dodged/Filled Bar Plot	▪ Chi-square test of Independence	▪ Observations are independent ▪ EXPECTED counts should be greater than 5

Variables of Interest	Parameter of Interest	Statistic of Interest	Descriptive Methods	Inferential Methods	Assumptions for Inferential Methods
Single Numerical Variable	True Population Mean (μ)	Sample Mean ($\bar{x}$)	▪ Report measures of center and variation ▪ Dotplot, boxplot, histogram ▪ Describe shape ▪ Identify outliers	▪ One-sample t-test ▪ CI for μ	▪ Either the sample size is fairly large or the data reasonably follow a normal distribution
Comparing Numerical Variable across Two Categories of a Categorical Variable (DEPENDENT samples)	True Mean Difference ($\mu_{difference}$)	Sample Mean Difference ($\bar{x}_{difference}$)	▪ Report measures of center and variation for the differences ▪ Dotplot, boxplot, histogram of the differences ▪ Describe shape, identify outliers	▪ paired t-test ▪ CI for $\mu_{difference}$	▪ Independent differences ▪ Either the number of pairs is fairly large or the differences reasonably follow a normal distribution
Comparing Numerical Variable across Two Categories of a Categorical Variable (INDEPENDENT samples)	Difference in True Population Means ($\mu_1 - \mu_2$)	Difference in Sample Means ($\bar{x}_1 - \bar{x}_2$)	▪ Report $\bar{x}_1, \bar{x}_2,$ and s_1, s_2 ▪ Side-by-side boxplots, facetted histograms	▪ Two-sample independent t-test ▪ CI for $\mu_1 - \mu_2$	▪ Observations are independent *between* & *within* groups ▪ Either both sample sizes are fairly large or the data from each group reasonably follow a normal distribution

Variables of Interest	Parameter of Interest	Statistic of Interest	Descriptive Methods	Inferential Methods	Assumptions for Inferential Methods
Comparing Numerical Variable across 2 or more categories of a Categorical Variable			▪ Group means, group standard dev. ▪ Side-by-side boxplots, facetted histograms	Analysis of Variance (ANOVA) F-test statistic	▪ Independence *between* and *within* groups ▪ Equal variances ▪ Normality
Comparing Two Numerical Variables	Population Slope (β_1)	Sample Slope (b_1)	▪ Correlation (r) ▪ Scatterplot ▪ Regression line ($\hat{y} = b_0 + b_1 x$)	Linear Regression Analysis Slope = 0?	▪ Linearity ▪ Independence ▪ Normality ▪ Equal Variance