|  |
| --- |
| **Comparing Sample Sizes for Each Group of a Categorical Variable** |
| **Creating a Summary Table of Observations of One Variable**  <NAME OF DATASET> %>%  count(<NAME OF CATEGORICAL VARIABLE>) |
| **Creating a Summary Table of Observations from Two Variables**  <NAME OF DATASET> %>%  count(<NAME OF CATEGORICAL VARIABLE 1>,  <NAME OF CATEGORICAL VARIABLE 2>) |
| **Simulation-Based One-Way ANOVA for a Variable with THREE or More Levels** |
| **Obtaining the Sample (Observed) F-Statistic**  obs\_F <- <NAME OF DATASET> %>%  specify(response = <NAME OF VARIABLE>,  explanatory = <NAME OF VARIABLE>) %>%  calculate(stat = "F") |
| **Obtaining 1000 Permuted F-Statistics**  null\_distribution <- <NAME OF DATASET> %>%  specify(response = <NAME OF NUMERICAL VARIABLE>,  explanatory = <NAME OF CATEGORICAL VARIABLE>) %>%  hypothesize(null = “independence”) %>%  generate(reps = 1000, type = "permute") %>%  calculate(stat = "F") |
| **Plotting the Null Distribution**  visualize(null\_distribution) +  labs(x = “Permuted F-statistic”) |
| **Shading the p-value**  visualize(null\_distribution)+  labs(x = “Permuted F-statistic”) +  shade\_p\_value(obs\_stat = obs\_F,  direction = “greater”) |
| **Obtaining a p-value**  get\_p\_value(null\_distribution,  obs\_stat = obs\_F,  direction = “greater”) |
| **Simulation-Based One-Way ANOVA for a Variable with TWO Levels** |
| **Obtaining the Sample (Observed) Difference in Means**  obs\_diff <- <NAME OF DATASET> %>%  specify(response = <NAME OF VARIABLE>,  explanatory = <NAME OF VARIABLE>) %>%  calculate(stat = "diff in means") |
| **Obtaining 1000 Permuted Differences in Means**  null\_distribution <- <NAME OF DATASET> %>%  specify(response = <NAME OF NUMERICAL VARIABLE>,  explanatory = <NAME OF CATEGORICAL VARIABLE>) %>%  hypothesize(null = “independence”) %>%  generate(reps = 1000, type = "permute") %>%  calculate(stat = "diff in means") |
| **Plotting the Null Distribution**  visualize(null\_distribution) +  labs(x = “Permuted Difference in Means (<GROUP 1> - <GROUP 2>)”)  ***Note:*** You should state what order of subtraction is being used for the difference in means! |
| **Shading the p-value**  visualize(null\_distribution)+  labs(x = “Permuted Difference in Means (<GROUP 1> - <GROUP 2>)”) +  shade\_p\_value(obs\_stat = obs\_diff,  direction = “two-sided”)  ***Note:*** You should state what order of subtraction is being used for the difference in means! |
| **Obtaining a p-value**  get\_p\_value(null\_distribution,  obs\_stat = obs\_diff,  direction = “two-sided”) |
| **Theory-based One-Way ANOVA** |
| **Conducting a Theory-based One-Way ANOVA**  my\_model <- lm(<NAME OF RESPONSE VARIABLE> ~ <NAME OF EXPLANATORY VARIABLE>, data = <NAME OF DATASET>)  anova(my\_model) |