Analysis Plan Document

Project Name: [Title]
Project Code: [Code]
Date Finalized: [Date]

This document serves as a basis for distinguishing between planned (confirmatory) analysis and any unplanned (exploratory) analysis that might be conducted on project data. This is crucial to ensuring that results of statistical tests will be properly interpreted and reported. In order that the Analysis Plan fulfill this purpose, it is essential that it be finalized and date-stamped before we begin looking at the data — ideally, before we take possession of the data. Once this plan is finalized, a date is entered above, and the document is posted publicly on our team website.

Data and Data Structure

This section describes variables that will be analyzed, as well as changes that will be made to the raw data with respect to data structure and variables.

Outcome Variables to Be Analyzed:
[List all outcome variables for which analysis is planned. Specify which outcomes are primary outcomes and which are secondary outcomes. Specify which outcomes are in the raw data and which outcomes will be created.]

Transformations of Variables:
[Describe new variables that will be created by transforming or combining variables in the raw data. If you plan on transforming, centering, or recoding the data, or will apply a coding scheme for categorical variables, please describe that process. If any measurements are going to be combined into an index (or even a mean), what measures will you use and how will they be combined? Include either a formula or a precise description of your method.]

Imported Variables:
[Describe variables that will be added to the raw data. These may be variables used in the experimental design (e.g., blocking variables) or covariates merged in from a different dataset.]

Transformations of Data Structure:
[Describe any plans to transform the structure of your data (e.g., from cross-sectional to panel).]

Data Exclusion:
[Describe criteria you will use to determine whether any observations should be excluded from analysis (e.g., values that are likely data-recording errors, outliers, etc.).]

Treatment of Missing Data:

[Describe your plan to handle missing data. How will you deal with incomplete or missing data (e.g., pairwise or listwise deletion, imputation, interpolation)?]

Statistical Models & Hypothesis Tests

This section describes the statistical models and hypothesis tests that will make up the analysis — including any follow-ups on effects in the main statistical model and any exploratory analyses that can be anticipated prior to analysis.

Statistical Models:

[Specify your main statistical model. What statistical model will you use to test each hypothesis? Please include the type of model and the specification of the model. This includes all variables that will be included as predictors, outcomes, or covariates. If you plan to test null hypotheses other than equality (zero effect), please specify the type of test and the margin. Explain the rationale for each covariate used. Please specify any interactions that will be tested and any planned checks on model assumptions. Remember that any test not included here must be noted as an exploratory test in any report of this research.]

Follow-Up Analyses:

[Specify your follow-up analyses. If not specified previously, will you be conducting any planned analyses to follow up on effects in your statistical model, such as subgroup analyses, pairwise or complex contrasts, or follow-up tests from interactions? For example, will you look for heterogeneous treatment effects if you find a main effect of the treatment? Will you look for specific pairwise differences between treatment arms if you find an overall effect of the pooled treatment? Note that these are planned analyses that will be conducted based on results of the main statistical model; exploratory analyses can be recorded in a later section.]

Inference Criteria, Including Any Adjustments for Multiple Comparisons:

[What criteria will you use to make inferences? Please describe the information you will use (e.g., specify the p-values), as well as cut-off criteria, where appropriate. Will you be using one- or two-tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, what method will you use to correct for multiple tests? If you are taking into account practical significance, please specify a minimum effect size of interest.]

Limitations:

[Describe any anticipated limitations associated with this analysis, if not previously identified in the Project Design document.]

Exploratory Analysis:

[OPTIONAL: If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or where there are multiple possible tests that you are going to

use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could then be tested using a pre-specified analysis at a later time.]

Link to an Analysis Code/Script:

[This OPTIONAL step is helpful in order to create a process that is completely transparent and increase the likelihood that your analysis can be replicated. We recommend that code include clear commenting that is interpretable without extensive prior knowledge of the project. We recommend that you run the code on a simulated dataset in order to check that it will run without errors.]