Stat
The ISI's Journal for the Rapid Dissemination of Statistics Research
(wileyonlinelibrary.com) DOI: 10.100X/stat.???

Crafting Statistical Analysis Plans: a Cross-Discipline Approach

Kimberly A. Cressman[a,b*] and Julia L. Sharp[c]

Received 9/12/2022; Revised 11/1/2022; Accepted enter date

Developing a plan for data analysis at the beginning of a study is an important practice that is underutilized in many scientific fields. Several funding agencies and journals now require submission of statistical analysis plans in advance of scientific studies, particularly in the clinical sciences. Even when a plan is not required, it can be advantageous to the scientific process by improving reproducibility. An analysis plan allows researchers to organize their knowledge about their research questions and experimental design to more easily recognize and choose the appropriate statistical analyses. An analysis plan provides a roadmap for the analyses: researchers can think through potential statistical decisions (e.g., to transform or not to transform? how to handle missing or censored data?) in advance, and thoroughly document the justifications and tradeoffs for their intended analyses. Such decisions are not influenced by data when made before data collection, thus preventing pitfalls like p-hacking, HARKing, and non-replicability of results. We describe a general framework for crafting an analysis plan - including essential components of any plan - and provide an example template that can be used by researchers. The analysis plan framework is presented for broad appeal to experienced statisticians, quantitative researchers, and everyone in between.

Keywords: Collaboration, Communication, Documentation, Reproducibility, Statistical design, Study design

1. Introduction

While the methods section of publications tends to be short, many decisions lead up to the implementation of a study's data collection, analysis, and ultimately publication. These decisions include the study design and statistical analysis choices (e.g., to transform or not to transform? how to handle missing or censored data?). An analysis plan is a description of the steps of the analyses that will be used to understand study objectives (Yuan et al., 2019). The analysis plan is a part of the collaborative process (Pomann et al., 2021) and is a component of setting expectations and scope of work for a project. The analysis plan can also serve as documentation of the analysis, before and

[a]Grand Bay National Estuarine Research Reserve, Moss Point, MS, USA
[b]Mississippi State University, Biloxi, MS, USA
[c]Department of Statistics, Colorado State University, Fort Collins, CO, USA
*Email: cressmanka@gmail.com

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

after the study. The plan can help to organize thoughts and support time management and organization. The plan can also be used as a template for a final report or a methods section of a journal article.

Many papers have been written in recent years to help scientists organize their data for easier analysis and sharing (e.g., Broman & Woo, 2018; White et al., 2013). Other publications are meant to help these scientists understand and improve reporting on the statistical procedures they may be using (e.g., Bolker et al., 2009; Dwivedi & Shukla, 2019; Harrison et al., 2018; Pederson, Miller, Simpson, & Ross, 2019; Zuur, Ieno, & Elphick, 2010; Zuur & Ieno, 2016), and still others have broached the statistical debate about p-values for non-statisticians (Smith, 2020). Some funding agencies and journals have begun to require pre-registration of study designs and statistical analysis plans prior to conducting a study, in a push for transparency and reproducibility (Banks et al., 2019; Kimmelman, 2021; Munafo et al., 2017). Pre-registration allows for investigators to distinguish between confirmatory research and exploratory research (see, for example, the Center for Open Science pre-registration website; Munafo et al., 2017) which can be critically important for the identification of a priori and post hoc analyses (Banks et al., 2019). Yuan et al. (2019) and Simpson (2015) provide detailed guidance on analysis plan development for biomedical and social sciences, respectively, in the context of pre-registration. Developing an analysis plan even in contexts where pre-registration is not required provides researchers with similar benefits, especially if they are willing to share the plan with others. An analysis plan can help prevent the inclination to report only a portion of results that are statistically significant (p-hacking) and the practice of hypothesizing after the results are known (HARKing). At a minimum, constructing a plan will help researchers consider how to address these various issues.

Crafting and implementing a plan can be essential to the success of any project. Literature on writing analysis plans is predominantly focused on the biomedical sciences, with fewer papers providing guidance in other contexts, relevant to a broad range of disciplines (e.g., ecology, agricultural sciences, education, sociology). Yuan et al. (2019) noted 49 components of a statistical analysis plan for biomedical research including randomization, analysis methods, covariates and adjustments, and subgroup analysis. Templates have been created for statistical analysis plans in the clinical and translational sciences (see for example "Duke Biostatistics", n.d.). While many of these components extend to other disciplines, we broaden the scope of the analysis plan for generalizing to other fields. In sum, the analysis plan is a roadmap of the analyses to be conducted, similar in construct to that of a data management plan (Michener, 2015).

The plan should be clearly documented and the steps detailed should be reproducible. Peng (2015) describes that for a study to be reproducible, the raw data from the study should be made available "and that the statistical code and documentation to reproduce the analysis are also available." Ellis

and Leek (2018) add to this statement that at the publication phase of the study, "reviewers and the rest of the world should be able to exactly replicate the analyses from raw data all the way to final results."

It is important to think about statistics during the study design phase. Studies may have already been conducted to collect the data (retrospective), or the data collection may happen in the future (prospective). Regardless, the statistical analysis plan will provide a guide for conducting the analysis. Many texts across disciplines have at least one chapter on the importance of the research question and study design. For example, at least one chapter in texts focused on various types of ecological study, ranging from long-term environmental monitoring (Gitzen, Millspaugh, Cooper, & Licht, 2012), population ecology (MacKenzie et al., 2018; Murray & Sandercock, 2020), species censuses (Sutherland, 2006) and even overall ecological statistics (Gotelli & Ellison, 2018; Zuur, Ieno, & Smith, 2007) are dedicated to this topic. Navarro (n.d.) states that "statistics is deeply intertwined with research design," in their introductory statistics text for psychology students (p. 9). More generally, in their step-by-step grant writing book, O'Neal-McElrath (2013) stresses that the evaluation methods must be established before project implementation (p. 56). In many fields, a first course in research methods sets the stage for future statistical analyses. By considering statistics before the study is implemented (or before the statistical analyses are conducted), the study team can carefully detail the design of the study including understanding the appropriate number of samples or observations needed (prospective), the level of the units in the study relevant to replication (e.g., experimental versus observational unit; prospective or retrospective), and time, cost, and effort (prospective or retrospective). Makin and de Xivry (2019) state that, "defining the analysis criteria in advance and independently of the data will protect researchers from circular analysis," which includes choosing a subset of the data that is relevant to the statistical results of the study.

As with choices involved in study design, choices involved in statistical analyses often involve tradeoffs. It may be - and perhaps often is - the case that any differences to conclusions resulting from choices made are subtle. That is, multiple reasonable analyses, and associated choices, would not change the results of a study. There is almost never a single "correct" statistical analysis to run (Lavine, 2019 as quoted in Wasserstein, Schirm, & Lazar, 2019; Munafo et al., 2017; Wagenmakers, Sarafoglou, & Aczel, 2022) and researchers may even consider running multiple analyses (e.g., sensitivity analysis) to examine whether and how results differ (Helsel, Hirsch, Ryberg, Archfield, & Gilroy, 2020). Ultimately, the statistical choices will need to be justified in final reports and publications. The analysis plan gives the research team space and context in which to document these decision points, potential trade-offs, and the final choices, in much more detail than would be included in the methods section of a paper.

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

The analysis plan might be written by the person implementing the steps in the plan, in collaboration with this person, or for this person (henceforth 'the statistician'). Ideally, the person implementing the steps in the plan would be involved in setting the plan by collaborating with the project lead and team. Researchers can provide detail on the research question, planned study design, a data dictionary, and anticipated outcomes (data ranges and variable distributions; hypothesized differences). The statistician can synthesize background information on the project and suggest appropriate analyses - and, when needed, changes to the study design (if, e.g., sub-samples are treated as independent replicates). Even those researchers that may be shy about statistics (or especially those researchers) will benefit from articulating the details of study design in a statistical analysis plan, as the plan organizes study knowledge in a way to more easily recognize, choose, and ultimately apply appropriate statistical analyses. When the analysis plan is constructed, the entire collaboration team is the audience for the plan. Everyone on the team should be comfortable with how the data will be statistically analyzed to answer the research questions. As the plan evolves, the audience may be a smaller subset of the team; for example, the statistician and the team leaders may be the individuals with detailed knowledge of the analysis plan at the implementation stage. The analysis plan may additionally be shared with future team members, and even external researchers and/or statisticians after the project is over. External audiences are not the primary audience for these plans, so authors should focus their time on capturing the ideas and details rather than polishing the writing.

An analysis plan is generally more detailed than a scope of work document. A scope of work (SOW) is an agreement between two parties that might include broad strokes of the work to be done. Specifically, "a SOW is a document drafted by a statistical consultant…which defines a research problem in statistical terms and specifies a set of deliverables, conditions, and expectations" (Peterson et al., 2022). An analysis plan would include substantial detail on aspects of the project that may only need to be known by the person completing the work. For example, in a scope of work, we might state 'Conduct a mixed effects model analysis' whereas in an analysis plan, we would detail the variables, the type of variables (e.g., fixed, random, response, predictor), the expected results, possible alternatives to the analyses, and challenges that may occur. A scope of work would also include an estimate of the time and costs to complete the goals of the project.

In this manuscript, we describe a general framework for crafting an analysis plan, while also detailing essential components of any plan. We provide a table to support translating questions about the scientific domain and study into components of the analysis plan. Finally, we present a sample plan in the context of ecological sciences. The analysis plan framework is presented for broad appeal to experienced statisticians, quantitative researchers, and everyone in between. This framework can also be used in both observational studies and experiments.

2. Crafting the Analysis Plan

The analysis plan format and content will be unique to the statistician, the collaborator, or the scientific domain. Generally, there are seven key questions that should be considered in any analysis plan:

1) What is/are the research questions or objectives?
2) What are the variables that will be measured in the study, and what will the data for each look like?
3) What is the study design, and why is the study set up this way?
4) How will data and any patterns/results be summarized?
5) What will be examined statistically?
6) What alternative strategies should be considered, and when?
7) What are the statistical results that will be presented?

These questions are related to and expand upon Chatfield's (1995) "stages of an idealized statistical investigation." More detail regarding these questions is shown in Table 1 and in Sections 2.1-2.7.

2.1. Research Questions/Objectives

The first consideration in crafting an analysis plan is to establish measurable objectives relevant to the study design and the variables collected or measured. Doran, Miller, and Cunningham (1981) coined an acronym, SMART, for writing meaningful objectives: Specific to an area of improvement; Measurable to quantify or indicate progress; Assignable to someone that will implement the work; ensuring that the results will be Realistic within the scope of the project constraints; and Time-related to establish a timeframe for the study. While several adaptations of the acronym have been made, the underlying theme is that objectives should be well-defined and able to be assessed for progress. Dwivedi (2022) states that "[c]lear reporting of data analyses as per the study objective…minimize[s] heterogeneous practices and improve[s] scientific quality and outcomes." In terms of the analysis plan, the research questions or objectives should be tied to the study design, variables, and data collected. A first draft of the objectives may need to be refined as other portions of the analysis plan are refined.

2.2. Variables and Their Types

Stat
The ISI's Journal for the Rapid Dissemination of Statistics Research
(wileyonlinelibrary.com) DOI: 10.100X/stat.???

It is important to carefully define the relevant variables in relation to study objectives to ensure that we can make appropriate conclusions from the statistical analyses. Gelman and Loken (2014) provide an example of how essential the choice of variable definition can be in the interpretation of statistical results and potentially providing an answer to the wrong question. They further discuss how a single scientific hypothesis can lead to many statistical hypotheses, and this can inadvertently occur without conscientiously defining variables before the study is implemented (Gelman & Loken, 2014). In this regard, once objectives have been drafted, the variables that will be measured or collected - and the units of each - should be specifically defined as relevant to the objectives. For example, in a study where the investigators would like to examine changes within individuals over time, they would need to collect data on the same individuals at multiple time points rather than multiple groups of individuals. The variables that we would formulate would be time and individual, rather than time, group, and individual. Even when data have already been collected, it is useful to define the variables and the units to ensure that those variables are appropriate for the stated objectives. For some, defining the variables while considering the eventual statistical model that will be used can be helpful.

The variables should also be classified as dependent (other example terms for this type of variable: response, outcome) or independent (other terms for this type of variable: explanatory, predictor, covariate, control, mediator, moderator) variables. Along with the definition of the variable, the variable type should also be described in the plan: quantitative (i.e., discrete, continuous) or qualitative (i.e., ordinal, nominal). For some variables, like age, the description may be related to the context. For example, in one study, researchers may collect ages of individuals using categories (e.g., 0-12 months, 12-24 months), while in other studies ages of individuals may be collected using a specific value (e.g., 8 months). Still another study may collect the ages of individuals using a value and then convert them to categories. The description of the variables should also include a general range of anticipated values. In some fields, zeros, missing data, below detection data, and/or censored data may be recorded. It is important to clearly describe each of these so that the data can be used (or not) appropriately in the analyses. For example, in some analyses, software will exclude observations when there is missing data. If missing observations are replaced with zeros, the estimates of the parameters will be biased downward.

2.3. Study Design

The study design is also important when developing a statistical analysis plan. A major component of the study design portion of the plan is consideration of the sample size. In some studies, the sample size is calculated based on a priori information like effect size, significance level, and anticipated power. In other studies, sample size may be fixed without the possibility of change. Sample size is

often the result of feasibility and practicality (e.g., cost and time). Regardless, the sample size for a study should be carefully appraised, motivated, and described in the study design section of the analysis plan, even when the data have already been collected.

In experimental studies, the study design is sometimes thought of as the 'design or blocking structure' and the 'treatment structure' (see, e.g., Stroup, 2012). The 'design or blocking structure' defines the method of controlling variation; this could be in the form of blocks, nesting, or repeated measurements on the same individual. The 'treatment structure' is the way in which the treatments are applied or organized in an experiment or how groups are structured; this could be one factor or group, two factors or groups, etc. Both the design and the treatment structures are related to the experimental units (i.e., replicates) and the observational units. Experimental units are the objects or people for which a treatment is applied, and are considered the independent units of replication in a study. Observational units are the objects or people for which a measurement or observation is made. *Experimental units and observational units are not necessarily the same.* For example, in a study of the impacts of a fertilizer on plants, suppose two fertilizer levels are each applied to one bench and then the growth of the plants on the bench is measured. The experimental unit here is the bench while the observational unit is the plant. To achieve independent replicates of the plants, assignment of each of the fertilizers on each of the plants would need to be randomized. While this framing is often used for experimental studies, it can be extended to observational studies as well.

Although different fields may have different terminology, similar ideas are applicable. For example, in ecology, a researcher might measure several bird chicks from the same clutch, and several clutches from different females (Harrison et al., 2018). The variable measured - growth rate, for example - is likely to be similar within clutches. Observations within each clutch are not independent and are often referred to as pseudo-replicates. As another example, consider a study of culture medium to estimate cell growth. Ten batches of the medium will be prepared, and each batch will be used on three separate plates (30 plates in total). The observations from the plates within a batch are correlated. Batches are known as biological replicates and the plates within the batches are known as technical replicates (see Lazic, 2010 for more examples). Additionally, Gelman and Hill (2007) describe a study to estimate the distribution of radon levels by county in the United States. The two levels of hierarchy for a potential hierarchical linear model described were houses (level 1) and counties (level 2), where radon levels within a county might be correlated.

Treating observations as if they were independent replicates could lead to a different than expected Type I error rate in significance tests, along with misleading confidence intervals. In many of these cases, Type I error will be too high, and confidence intervals will be too narrow (Harrison et al., 2018; Hurlbert, 1984). Terms for designs like this are 'hierarchical' and/or 'nested'; in discussing these designs in terms of statistical analyses and models, the terms 'levels' or 'fixed' and 'random' effects

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

are used. While all of this vocabulary may not be in a researcher's toolbox (e.g., the idea of random effects can be particularly confusing), researchers often understand the concepts (e.g., researchers know to be concerned about non-independence of data). When a researcher constructs an analysis plan where a sampling structure or study design is described as we propose, with sufficient context and detail, the statistician will be able to suggest an appropriate analysis (e.g., mixed models).

2.4. Data Summary

The first components of the statistical analysis plan are essential for contributing to the subsequent parts and to understanding the data. The statistical analysis plan should include details on the particular and appropriate descriptive statistics that will be computed for each of the variables in the study. These would include measures of location (e.g., mean, median, quantiles) and measures of variation (e.g., standard deviation, inter-quartile range, coefficient of variation) for quantitative variables and frequencies and proportions for qualitative variables. Certain data quality checks, such as methods used to identify outliers and missing data patterns, can also be detailed in the data summary. For example, outliers may be identified using Tukey's 1.5 x IQR fence rule (Tukey, 1977). At this point, the types of figures that will be used to illustrate the data may also be considered. We have found that creating mock tables and drafting the figures on paper are helpful when it is time to implement the plan; these aspects are not essential in the plan itself, but are helpful to think through.

2.5. Statistical Analysis

The methods of statistical examination are the next elements to include in the statistical analysis plan. This section of the analysis plan is where discussion of the various analytical choices and tradeoffs of those choices can occur. While the following is not an exhaustive list, major topics that may be included in this section are: the type of statistical examination that will be conducted; what each analysis will demonstrate about the data (e.g., describe a relationship, predict new values, infer differences between groups); why a given analysis has been chosen over other possibilities (e.g., use of a non-parametric rather than parametric method, use of one transformation over another possibility); how the design and treatment features of the study contribute to modeling decisions; what post-hoc analyses may occur, and what drawbacks or cautions should be considered when interpreting results. Citing sources is particularly helpful in this portion of the plan - it will inevitably be necessary in the future to recall the source of certain details.

In this section of the plan, the researcher should reflect on all of the prior details (i.e., information from Sections 2.1-2.4) and use them to formulate an appropriate statistical analysis approach. For 'typical' analyses, researchers may refer to an analysis decision tree to help initiate discussion on the appropriate analysis to employ (with the understanding that the treatment and design structure may lead to more complex analyses; see for example Creaser, n.d.; Dwivedi, 2022; Nayak & Hazra, 2011; Simpson, 2015). In the research team's assessment of the appropriate analyses, they should consider the research questions, objectives, and aims and connect them with the variables that have been collected or measured to be included in the analyses (Gelman & Loken, 2014). These variables may include information about groups that will be compared, exploration of relationships among predictors and a response or multiple responses, factor identification, or community analysis.

2.6. Alternative Strategies

Several circumstances exist where alternative analysis strategies may need to be employed. In the alternative strategies section of the analysis plan, research teams should consider what to do if the analysis conducted does not go as expected (e.g., assumptions are not satisfied) or if they want to examine the model with different components (e.g., various covariance structures). For example, when assumption checks reveal that the assumptions of the analysis or the model are not satisfied, a transformation (e.g., a log transformation of the response variable) may be necessary. Or, if the model fit is not adequate, a different structure of the model could be considered. For example, if the residual plots show a parabolic shape, then a quadratic predictor may be included. Alternatively, if there are repeated measurements on the same individual, then different correlation structures might be examined. Uncertainty and sensitivity analyses (e.g., Saltelli et al., 2008) may be employed to examine how model results differ when model parameters are changed. We stress that these alternative strategies and in which situations they would be used should be planned in advance, as researchers should avoid making such decisions based on the data or the analysis results (Gelman & Loken, 2014; Makin & de Xivry, 2019; Munafo et al., 2017).

Missing data can also challenge the original, planned analyses; a plan for how to handle missing observations should be detailed. Whether missing data will be imputed or not should be noted, and the methods for imputation should be described. If the model can be used with missing data, then a statement in this section can help remind those implementing the analysis plan that nothing special must be done to handle the missing observations. When data collection results in many missing observations and imputation or other methods for accommodating missing data are not available, an appropriate strategy like providing descriptive statistics instead of inferential statistics could be detailed.

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

In some settings, rather than missing observations, there may be many observations that are zeros. Potential strategies in these cases are use of a Poisson or negative binomial distribution for modeling the variable, using a zero-inflated model, a hurdle model (a two-part model for the zeros and positive values separately), an ordinal regression model, or a descriptive analysis with a qualitative interpretation.

2.7. Presentation of Results

This portion of the analysis plan will first describe the presentation of data summary results (Section 2.4). From the statistical analysis results (Section 2.5), information to highlight from statistical analysis program output should be identified. We urge researchers to consult existing resources that discuss what to present from their analyses (e.g., Gotelli & Ellison, 2018 and Harrison et al., 2018 for mixed models; McCune, Grace, & Urban, 2002 for ecological community analyses; Zuur & Ieno, 2016 for regression analyses). Following the advice of Wasserstein et al. (2019) to accept uncertainty, be thoughtful, be open, and be modest, we recommend including a description of the estimates and their measures of uncertainty as well as other summary statistics that may benefit the interpretation of results (e.g., sample size, goodness of fit statistics). P-values are one summary statistic that if included should be presented as an equality (e.g., $p = .014$) and not as an inequality (unless the p-value is very small, e.g., $p < .001$), and interpreted correctly and conscientiously (Smith, 2020; Greenland, 2019 as quoted in Wasserstein et al., 2019).

3. Fish Monitoring Example Case Study

In this case study example, we tie the content of Table 1 to a long-term fish monitoring project in the Grand Bay National Estuarine Research Reserve (NERR), southeast Mississippi. NERR staff, working with local researchers, designed the study to identify fish species present in the NERR and to determine whether fish community varied by habitat type within the NERR (unpublished data). Sampling was conducted quarterly from 2005-2014. Ecological community datasets such as this can be complex, and like many, this one included environmental data (e.g., water temperature, salinity, dissolved oxygen, and tide stage) and species data.

We faced many decisions leading up to the data analysis, not least of which was "what statistical test should we use?" Knowing the characteristics of our data helped direct our focus when reviewing the literature by allowing us to key in on details about similar study designs and data collection methods. A Bray-Curtis dissimilarity matrix is commonly used in multivariate ecological analyses such as this study (Gotelli & Ellison, 2018). Decisions we did have to make, however, were: should the data be

transformed, and if so, how?; should all species be retained in the dataset or should the focus be on the most common species (and how would "most common" be defined)?; how should seasonality in the data be handled? We also had to consider the dependence in the dataset: because fish naturally move around and environmental conditions vary, the population available to be captured was not the same from season to season, or even necessarily in the same season across different years. Fish communities sampled on the same day could be more likely to resemble each other than communities sampled at the same site at a different time. This study is the basis for an example analysis plan (Sections 3.1-3.7) using the framework presented in Table 1. Some details have been condensed, adapted, or omitted to allow for clarity and completeness of the example. In actuality, much of this plan was written in outline form, rather than in clean, publishable syntax (see Figure 1). For some research teams' purposes, this approach may be adequate, though transitioning the language into that of a publication could require more work.

3.1. Fish Monitoring Research Questions/Objectives

Grand Bay NERR was established in 1999. As of 2005, there had not been comprehensive sampling to determine the composition of the fish community in the NERR. One purpose of establishing the fish monitoring project was to compile a baseline species list, including multiple habitat types and multiple seasons in order to capture spatial and temporal variation. Another objective of the project was to determine whether, and how, species composition differed among several key habitat types within the NERR.

3.2. Fish Monitoring Variables and Their Types

In this study, the fish community contains multiple species (categories). The species community data are represented by a matrix of the number of individuals of each species caught in each sample (see Figure 2, for example). Rows in the dataset will represent the 500+ individual sampling events and columns will contain individual species names (approximately 100). Values in each cell will be counts of each species caught in each sample. As with other ecological community datasets, there will be many zeros in the data, as well as several high species counts (in the thousands); though the majority of non-zero counts are expected to fall within the range "one to a few dozen".

Potential explanatory variables describe the sites (habitat type) or environmental conditions during sampling: date, season, water temperature (degrees C), air temperature (degrees C), dissolved oxygen level (mg/L), salinity (ppt), and tide stage (high/low, incoming/outgoing). These environmental variables will be measured directly in the field. Water and air temperature are

Stat
The ISI's Journal for the Rapid Dissemination of Statistics Research
(wileyonlinelibrary.com) DOI: 10.100X/stat.???

continuous variables that potentially follow a normal distribution. Both are likely to range from ~4 - 35 degrees C; outliers would be surprising. Dissolved oxygen and salinity are also continuous variables, but will probably not be normally distributed. The expected range of values for dissolved oxygen is ~2 - 8 mg/L. High or low outliers are both possible during abnormal events like algae blooms. Salinity values are expected to be between ~10 - 30 ppt. Low outliers are possible after extreme rain events and high outliers are possible during drought; even these would only expand the range to 0 - 35. Zero is a lower bound for salinity and would be a surprising value at any of these sites. All other non-date variables are categorical. We expect both ordering and cycling in season and tide stage, and no ordering (nominal only) in habitat type.

3.3. Fish Monitoring Study Design

Sampling will be conducted quarterly to capture all seasons. Fourteen sites will be selected at the beginning of the study, spanning five habitat types: beach (2 sites), shell midden (3 sites), erosional marsh edge (3 sites), depositional marsh edge (3 sites), and seagrass beds (3 sites). Sites are replicates within habitat type. The sites will be close enough to each other that location alone is not expected to cause differences in the fish community between sites (see Figure 17 in GBNERR, 2013). Fish will be captured by pulling a 6 m bag seine (a type of net that reaches from the bottom of the body of water to the surface) parallel to shore for 50 m, except for seagrass sites, where the seine will be pulled toward shore from 50 m away.

We need to consider potential imbalance in the study design. Not all habitat type categories contain the same number of sites: depositional marsh edge, erosional marsh edge, shell midden, and seagrass habitats each have 3 sites; while beach only has 2. Additionally, we need to ensure that water quality parameters (temperature, salinity, dissolved oxygen) are not so dissimilar across sites that fish community differences could result from differences in water quality characteristics rather than differences in habitat (i.e., confounding between habitat and water quality). Seasonality is expected in the data, both environmentally and biologically. For this analysis, we need to account for seasonality but it is only tangential to the inferential goals.

3.4. Fish Monitoring Data Summary

During exploratory data analysis, we will use graphical displays to ensure salinity and dissolved oxygen (response variables) are similar across sites. Box plots will be made for each of the response variables categorized by site (horizontal axis) to examine overall conditions. To explore conditions through time, line graphs with date along the horizontal axis, response variable values on the y-axis,

and sites denoted by unique combinations of color and shape will be employed. If any sites appear different in biologically important ways (e.g., consistently lower salinity), those sites will be dropped from further analyses.

We will construct a table to summarize water quality data, presenting seasonal mean, median, and standard deviation for each parameter, for each habitat type. We will group sites within each habitat so the table is a reasonable size.

Tables will also be used to summarize fish community: 1) Total counts (over all 10 years) of most abundant species. All species will be summarized for the final report supplementary information; 2) Counts of most common species, by habitat; 3) Counts of most common species, by season.

Figures generated for reporting will be plots resulting from non-metric multidimensional scaling (NMDS). Vectors for species that are identified as important from SIMPER will be overlaid on the NMDS plots.

3.5. Fish Monitoring Statistical Analysis

To reduce the impact of dominant species on dissimilarity calculations without potentially over-weighting rare species, all species counts will be square-root transformed. A log transformation may also be considered, but we want "medium-abundance species" to be adequately represented (Clarke & Green, 1988; McCune et al., 2002). Many community studies only retain the most common species for analysis (Clarke & Green, 1988), though this is not ideal for hypothesis testing (McCune et al., 2002). We will thus retain all species in the statistical analyses.

To describe community differences between habitat types, we will use Analysis of Similarities (ANOSIM; Clarke, 1993; Field, Clarke, & Warwick 1982). ANOSIM output consists of a measure of dissimilarity, known as 'R', and a p-value. R typically ranges from 0-1, with higher values indicating more of a difference between groups. The p-value is calculated via permutations and is a measure of how extreme the R value is, compared to all possible R values for the permuted sample data. Due to the likelihood of temporal variation in the area's fish population, we need to condition on the population available at the time of sampling; permutations will thus be restricted to block on sampling date.

It is possible that season can impact community differences between habitats. ANOSIM cannot model interactions (Somerfield, Clarke, & Gorley, 2021) so we will perform statistical analyses on each season separately. If differences between habitat groups are observed from the global ANOSIM for a season ($p < 0.05$), we will consider this as sufficient evidence that at least one group is different

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

from at least one other. In this case, we will run ANOSIMs between each pair of habitats to examine pairwise differences. Because this analysis is exploratory, we will not adjust p-values on these pairwise comparisons.

When pairwise differences are significant in ANOSIM, the Similarity of Percentages (SIMPER; Clarke, 1993) method will be employed to identify which species contribute most to those differences. Output from the *simper* function in the R statistical software package *vegan* (Oksanen et al., 2022) will list the species that most contribute to community differences between the pair of groups, as well as the average contribution of "each species to the average overall Bray-Curtis dissimilarity" (Oksanen et al., 2022). SIMPER is known to place undue weight on species that are more variable rather than necessarily distinctive (Warton, Wright, & Wang, 2012), so these results must be interpreted with caution.

Non-metric multidimensional scaling (NMDS; Clarke, 1993) will be used to visualize differences between habitat types. As with ANOSIM, NMDS will be run on each season individually. Scree plots of NMDS stress values will be used to determine how many axes should be used in NMDS.

3.6. Fish Monitoring Alternative Strategies

We plan to use ANOSIM because the data are not typically normally distributed, and the design is unbalanced (with some habitats having 3 site replicates and others having only 2). We may also explore PERMANOVA but note that it is not robust to imbalance as is ANOSIM (Walters & Coen, 2006). We could also investigate the analysis of log transformed data. The log transformation gives more weight to less common species than a square-root transformation (Clarke & Green, 1988; McCune et al., 2002) and may slightly change which species are identified as important contributors to dissimilarity between habitats. While these differences may be subtle, they could also be informative.

3.7. Fish Monitoring Presentation of Results

We will combine statistical output into a single table, if possible, as the goal is to use all methods to understand fish community variability by habitat type. A draft table, to be filled in with specifics, is presented as Figure 3. The table will include ANOSIM's R and p-values and the NMDS stress value. Each pairwise comparison for each season will have a row for the pairwise ANOSIM R and p-values, and if the ANOSIM indicates a significant difference between the pair of habitats ($p < .05$), the top three species that SIMPER identifies as contributing to the dissimilarity will be listed.

4. Discussion

We emphasize that the analysis plan should be considered in conjunction with the study design. The analysis plan does not need to be completely cemented before data are collected, as changes to the design and data collection may occur during study implementation. However, if a plan is drafted before data are collected (prospective), and changes do result during implementation, the statistician can be involved on the team to discuss these changes and participate in the conversation about how the changes may impact the analysis plan. If data have already been collected (retrospective), the statistician can help to ensure that the variable definition and units are relevant to the objectives of the analysis to ultimately support answering the research questions. Changes can be documented in the plan, providing transparency. The analysis plan should be a living document, like data management plans (Michener, 2015). Changes in the analysis plan could be documented in a 'revision history' section (Yuan et al., 2019) or using version control software (e.g., Git, Google Docs). An additional benefit of a well-crafted analysis plan is that the methods section of a report or manuscript could be framed similarly to the plan, potentially changing future tense from the plan to past tense in the methods section.

Researchers and statisticians, separately and together, will benefit from crafting an analysis plan. Researchers may feel stymied by creating an analysis plan on their own based on previous experiences with statistics. By walking through an analysis plan as a series of steps and questions rooted first in the research question and variables, rather than worrying about exact statistical terminology, researchers themselves can better understand and justify the decisions that need to be made throughout the analysis process. We stress that the 'alternative strategies' section of the analysis plan is not meant to promote p-hacking or HARKing. Rather, recognizing that unforeseen challenges may occur during the study implementation and subsequent analysis, we advocate for thinking through potential challenges during the analysis plan stage of the research process. The 'alternative strategies' section provides a framework for identifying and documenting strategies to overcome these challenges, allowing the research team to thoughtfully plan and promote transparency in research. For example, for the fish monitoring case study, writing out the analysis plan helped the research team get on the same page about what data were in hand, how all samples and variables related to each other, and what the problems and caveats might be with our desired analyses. Articulating the details has contributed to writing of the methods section of the fish monitoring study report (unpublished) and provided a useful log of our analysis pathway. Documentation of our analysis decisions also allowed for statistical reproducibility, or even methodology comparisons, should future researchers wish to see if different statistical choices (e.g., using a log- rather than square-root transformation; excluding rare species) impact the conclusions.

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

Talking with a statistician about the study design early in the process, and especially providing the details the researcher *is* knowledgeable about at that point, can help the researcher think about the analysis plan not as a statistical hurdle or qualitative project, but as a stepwise process to understanding the story that the data tell. While it is possible to come up with a plan that seems reasonable, talking through the details of the study design and analysis with others can help to ensure the study's integrity.

The analysis plan approach that we consider here focuses predominantly on quantitative analyses. The plan can be extended for qualitative and mixed methods (i.e., quantitative and qualitative) studies by asking and answering similar questions to those in Table 1. The approach presented here can be generalized to many fields and can be extended or honed to be field-specific.

The analysis plan is a pertinent component to a scope of work and has been referred to as a 'sketch' in this context (Peterson et al., 2022). The scope of work may include less detail than the analysis plan presented here. The time and effort typically involved in creating an analysis plan can vary depending on factors like study type, objectives, length, role on the project, experience, and complexity of the analyses. Some statisticians may be interested in composing one document that includes both the scope of work with a sketch of the analysis plan as well as the detailed analysis plan presented here. Future work in this area could include creating template documentation that renders portions relevant to each component.

Successful, impactful, and rigorous research is built on a solid foundation. Detailing study design, variables, analysis choices, and alternative strategies in an analysis plan allows for reproducibility, transparency, and documentation of the eventual statistical analyses, and is thus a meaningful contribution to this foundation. Researchers that are comfortable with statistics may be able to draft an analysis plan on their own, while those who are less comfortable with statistics can use our template to document study details that a statistician will need to know in order to assist. We encourage researchers and statisticians to work together on these plans, as leveraging their complementary expertise will strengthen the totality of the study and contribute to fruitful collaborations in the future.

Acknowledgements

The authors thank Jonathan Pitchford, Eric Sparks, and Brianna Andrews for feedback on draft versions of this analysis plan template. We thank two anonymous reviewers for their helpful comments on the manuscript. We also thank all Grand Bay NERR and Mississippi Department of Marine Resources (MDMR) staff that worked on the fish monitoring project, especially: Gretchen Grammer, Jake Walker, Brenna Ehmen, Ron Cole, Cher Griffin, Michael Brochard, Ayesha Gray, and

Paul Mickle. The National Oceanographic and Atmospheric Administration (NOAA) and MDMR provided funding for the fish monitoring project, and KC's time writing this manuscript was partly supported by NOAA grant NA22NOS4200071.

Bibliography

Banks, G. C., Field, J. G., Oswald, F. L., O'Boyle, E. H., Landis, R. S., Rupp, D. E., & Rogelberg, S. G. (2019). Answers to 18 Questions About Open Science Practices. *Journal of Business and Psychology*, *34*(3), 257–270. https://doi.org/10.1007/s10869-018-9547-8

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Broman, K. W., & Woo, K. H. (2018). Data Organization in Spreadsheets. *The American Statistician*, *72*(1), 2–10. https://doi.org/10.1080/00031305.2017.1375989

Center for Open Science (n.d.), "What is Preregistration?," Accessed October 28, 2022. Retrieved from https://www.cos.io/initiatives/prereg.

Chatfield, C. (1995). *Problem Solving: A statistician's guide. 2nd edition*. Boca Raton, FL: Chapman & Hall/CRC Press.

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Austral Ecology*, *18*(1), 117–143. https://doi.org/10.1111/j.1442-9993.1993.tb00438.x

Clarke, K., & Green, R. (1988). Statistical design and analysis for a "biological effects" study. *Marine Ecology Progress Series*, *46*, 213–226. https://doi.org/10.3354/meps046213

Creaser, C. (n.d.). "How to choose the right statistical technique," Emerald Publishing How to guides: data & analysis. Accessed August 17, 2022. Retrieved from https://www.emeraldgrouppublishing.com/how-to/research/data-analysis/choose-right-statistical-technique

Doran, G. T., Miller, A., & Cunningham, J. (1981). There's a S.M.A.R.T. Way to Write Management's Goals and Objectives. *Management Review*, 70(11), 35-36.

Duke Biostatistics, Epidemiology, and Research Design Core. Statistical analysis plan (SAP) template. Accessed September 9, 2022. Retrieved from https://biostat.duke.edu/berd-methods-core.

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

Dwivedi, A. K. (2022). How to write statistical analysis section in medical research. *Journal of Investigative Medicine*, jim-2022-002479. https://doi.org/10.1136/jim-2022-002479

Dwivedi, A. K., & Shukla, R. (2019). Evidence-based statistical analysis and methods in biomedical research (SAMBR) checklists according to design features. *Cancer Reports*, *3*(4), e1211. https://doi.org/10.1002/cnr2.1211

Ellis, S. E., & Leek, J. T. (2018). How to Share Data for Collaboration. *The American Statistician*, *72*(1), 53–57. https://doi.org/10.1080/00031305.2017.1375987

Field, J. G., Clarke, K. R., & Warwick, R. M. (1982). A practical strategy for analysing multispecies distribution patterns. *Marine Ecology Progress Series, 8,* 37-52.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, United Kingdom: Cambridge University Press.

Gelman, A., & Loken, E. (2014). The statistical crisis in science: data-dependent analysis--a "garden of forking paths"--explains why many statistically significant comparisons don't hold up. *Scientific American*, *102*(6), 460.

Gitzen, R. A., Millspaugh, J. J., Cooper, A. B., & Licht, D. S. (Eds.). (2012). *Design and Analysis of Long-term Ecological Monitoring Studies*. Cambridge, United Kingdom: Cambridge University Press.

Gotelli, N. J., & Ellison, A. M. (2018). *A Primer of Ecological Statistics* (2nd ed.). Sunderland, MA: Sinauer Associates.

Grand Bay National Estuarine Research Reserve (GBNERR). (2013). *Grand Bay National Estuarine Research Reserve Management Plan 2013-2018*. Moss Point, MS: Grand Bay National Estuarine Research Reserve, Mississippi Department of Marine Resources. https://coast.noaa.gov/data/docs/nerrs/Reserves_GRD_MgmtPlan.pdf

Greenland, S. (2019). Valid p-values behave exactly as they should: Some misleading criticisms of p-values and their resolution with s-values. *The American Statistician*, *73*(sup1), 106-114.

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., … Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, *6*, e4794. https://doi.org/10.7717/peerj.4794

Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., & Gilroy, E.J. (2020). *Statistical methods in water resources. U.S. Geological Survey Techniques and Methods 4-A3*. Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/tm4a3

Hurlbert, S. H. (1984). Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs*, *54*(2), 187–211. https://doi.org/10.2307/1942661

Kimmelman, J. (2021). *Clinical Trials* to authors: Please pre-register your studies! *Clinical Trials*, *18*(6), 645–646. https://doi.org/10.1177/17407745211057186

Lavine, M. (2019) Frequentist, Bayes, or Other?, *The American Statistician*, 73:sup1, 312-318, https://doi.org/10.1080/00031305.2018.1459317

Lazic, S. E. (2010). The problem of pseudoreplication in neuroscientific studies: is it affecting your analysis?. *BMC neuroscience*, *11*(1), 1-17.

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2018). *Occupancy Estimation and Modeling. Inferring Patterns and Dynamics of Species Occurrence* (2nd ed.). London, United Kingdom: Academic Press.

Makin, T. R., & Orban de Xivry, J. J. (2019). Ten common statistical mistakes to watch out for when writing or reviewing a manuscript. *ELife*, *8*, e48175. https://doi.org/10.7554/eLife.48175

McCune, B., Grace, J. B., & Urban, D. L. (2002). *Analysis of ecological communities* (Vol. 28). MjM software design Gleneden Beach, OR.

Michener, W. K. (2015). Ten Simple Rules for Creating a Good Data Management Plan. *PLOS Computational Biology*, *11*(10), e1004525. https://doi.org/10.1371/journal.pcbi.1004525

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., … Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Murray, D. L., & Sandercock, B. K. (Eds.). (2020). *Population Ecology in Practice*. Hoboken, NJ: Wiley and Sons.

Nayak, B., & Hazra, A. (2011). How to choose the right statistical test? *Indian Journal of Ophthalmology*, *59*(2), 85. https://doi.org/10.4103/0301-4738.77005

Navarro, D. (n.d.). *Learning Statistics with R*. 613. https://learningstatisticswithr.com/lsr-0.6.pdf Accessed 10/27/2022.

O'Neal-McElrath, T. (2013). *Writing grants step by step: The complete workbook for planning, developing, and writing successful proposals*. San Francisco, CA: Wiley and Sons.

(wileyonlinelibrary.com) DOI: 10.100X/stat.???

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., … Weedon, J. (2022). *vegan: Community Ecology Package*. https://CRAN.R-project.org/package=vegan

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876. https://doi.org/10.7717/peerj.6876

Peng, R. (2015). The reproducibility crisis in science: A statistical counterattack. *Significance*, *12*(3), 30–32. https://doi.org/10.1111/j.1740-9713.2015.00827.x

Peterson, R. A., Hochheimer, C. J., Grunwald, G. K., Johnson, R. L., Wood, C., & Sammel, M. (2022). Reaping what you SOW: guidelines and strategies for writing scopes of work statistical consulting. *Stat*, in press.

Pomann, G.-M., Boulware, L. E., Cayetano, S. M., Desai, M., Enders, F. T., Gallis, J. A., … Thomas, S. M. (2021). Methods for training collaborative biostatisticians. *Journal of Clinical and Translational Science*, *5*(1), e26. https://doi.org/10.1017/cts.2020.518

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., ... & Tarantola, S. (2008). *Global sensitivity analysis: the primer*. West Sussex, England: Wiley & Sons.

Simpson, S. H. (2015). Creating a Data Analysis Plan: What to Consider When Choosing Statistics for a Study. *The Canadian Journal of Hospital Pharmacy*, *68*(4), 311–317. https://doi.org/10.4212/cjhp.v68i4.1471

Smith, E. P. (2020). Ending Reliance on Statistical Significance Will Improve Environmental Inference and Communication. *Estuaries and Coasts*, *43*(1), 1–6. https://doi.org/10.1007/s12237-019-00679-y

Somerfield, P. J., Clarke, K. R., & Gorley, R. N. (2021). Analysis of similarities (ANOSIM) for 2-way layouts using a generalised ANOSIM statistic, with comparative notes on Permutational Multivariate Analysis of Variance (PERMANOVA). *Austral Ecology*, *46*(6), 911–926. https://doi.org/10.1111/aec.13059

Stroup, W. W. (2012). *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*. Boca Raton, FL: CRC Press.

Sutherland, W. J. (Ed.). (2006). *Ecological Census Techniques: A Handbook* (2nd ed.). Cambridge, United Kingdom: Cambridge University Press.

Tukey J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Wagenmakers, E. J., Sarafoglou, A., & Aczel, B. (2022). One statistical analysis must not rule them all. *Nature*, *605*(7910), 423–425. https://doi.org/10.1038/d41586-022-01332-8

Walters, K., & Coen, L. D. (2006). A comparison of statistical approaches to analyzing community convergence between natural and constructed oyster reefs. *Journal of Experimental Marine Biology and Ecology*, *330*(1), 81–95. https://doi.org/10.1016/j.jembe.2005.12.018

Warton, D. I., Wright, S. T., & Wang, Y. (2012). Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution*, *3*(1), 89–101. https://doi.org/10.1111/j.2041-210X.2011.00127.x

Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a World Beyond " $p < 0.05$." *The American Statistician*, *73*(sup1), 1–19. https://doi.org/10.1080/00031305.2019.1583913

White, E., Baldridge, E., Brym, Z., Locey, K., McGlinn, D., & Supp, S. (2013). Nine simple ways to make it easier to (re)use your data. *Ideas in Ecology and Evolution*, *6*(2). https://doi.org/10.4033/iee.2013.6b.6.f

Yuan, I., Topjian, A. A., Kurth, C. D., Kirschen, M. P., Ward, C. G., Zhang, B., & Mensinger, J. L. (2019). Guide to the statistical analysis plan. *Pediatric Anesthesia*, *29*(3), 237–242. https://doi.org/10.1111/pan.13576

Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems: Data exploration. *Methods in Ecology and Evolution*, *1*(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Zuur, A. F., Ieno, E. N., & Smith, G. M. (2007). *Analysing ecological data*. Springer.

Zuur, A. F., & Ieno, E. N. (2016). A protocol for conducting and presenting results of regression-type analyses. *Methods in Ecology and Evolution*, *7*(6), 636–645. https://doi.org/10.1111/2041-210X.12577

Table 1. Key questions for crafting an analysis plan.

1.	**What is/are the research question(s) or objectives?**

Stat
The ISI's Journal for the Rapid Dissemination of Statistics Research
(wileyonlinelibrary.com) DOI: 10.100X/stat.???

2. **What are the variables that will be measured in the study? For each variable, what will the data look like?**

    a. Dependent/response
    b. Independent/explanatory
    c. Are there other variables collected that might be considered in the analysis?

    For each:
    a. What type of data is the variable: quantitative, count, proportion, percentage, categories?
    b. How is the variable measured or defined (if calculated)?
    c. What are the units of measure?
    d. What general range of values is anticipated?
    e. Are many outliers anticipated?
    f. Are many values expected to be '0', censored, truncated, or missing?
    g. What would a histogram of quantitative values be expected to look like?

3. **What is the study design, and why is the study set up this way?** A diagram and/or map can help communicate this.

    a. What is sampled, measured, and recorded for each individual/object/station/unit (and why)?
    b. If there are repeated measurements on the same individual/object/station/unit, how often are they collected (and why)?
    c. What are the independent individuals/objects/stations/units (i.e., replicates)?
    d. How are the data points related to each other? Think about relationships in both space and time – would sites or subjects closer to each other be more similar than more distant ones? Is seasonality expected?
    e. Are there other sources of variation that might need to be considered (e.g., streamflow, precipitation, farm site, age, weight, education level, employment)?

4. **How will data and any patterns/results be summarized?**

    a. ***Summary statistics:*** perhaps mean, standard deviation, range, sample size, number of missing observations, etc. for each variable. Should these be calculated by group?
    b. ***Tables and Figures:*** What will the main tables and figures look like? What will they communicate about the data?

5. **What will be examined statistically?** Link directly back to the research questions. We provide a few questions below, but refer the reader to questions that may be relevant to the context of their field.

    a. Will there be a comparison among groups?
        i. How many groups are there that will be compared?
        ii. What are the groups? How are they defined (think about spatial and temporal relationships described above)?
        iii. How many samples are in each group?
        iv. What differences among the groups are expected?
        v. What procedure is typically used to test for this type of difference in this research field?
        vi. If there are more than two groups, what type of adjustment will be used to control for multiple comparisons?

    b. Will there be an examination of the relationship(s) or association(s) among the variables?
        i. Are the observations independent, or might they be related (in space and/or time)?
        ii. Is there more than one potential explanatory variable?
        iii. Is the interest in describing the direction and strength of a relationship, or in finding the best way to predict the response variable from the explanatory variable(s) in a new situation?
        iv. Is there interest in understanding the latent variable structure of several variables (e.g., factor analysis)?
        v. Are there other variables that could impact the relationship to be examined? How will these other variables be accounted for? Will these other variables be measured as part of the project?

    c. Will there be an investigation of 'communities' (i.e., groups of species found together)?
        i. How is community defined if represented in tabular form: what are the rows (samples), and what are the columns (species: counts/relative abundances/presence-absence)?
        ii. Will univariate statistics (e.g., richness and/or diversity metrics) be calculated?
        iii. What differences among the groups are expected, and on what scale (spatially, temporally)?
        iv. What type of community analysis will be used (e.g., ordination, PERMANOVA, discriminant analyses)?

Stat
The ISI's Journal for the Rapid Dissemination of Statistics Research
(wileyonlinelibrary.com) DOI: 10.100X/stat.???

6. **What alternative strategies should be considered, and when?**

    a. When will alternative strategies be employed?
    b. What will be done if the model is not a good fit, assumptions are not satisfied, or the model does not converge?
    c. How will missing data be handled?

7. **What are the statistical results that will be presented? What is commonly presented (or *should be* presented) for someone to understand the output of the intended statistical analyses?**

    a. See also questions 4 and 5, as the data summaries and inferential statistics will tie in with what is displayed here.
    b. This will be ***more than just a p-value*** and could belong in a table (e.g., sample size, degrees of freedom, test statistic, confidence intervals, slope and R^2 (for regression-type analyses), etc.).

Test: ANOSIM – rank-based (non-parametric) test, appropriate for non-normally distributed, multivariate data. Also okay for experimental design to be unbalanced (NFM design is unbalanced due to different numbers of sites in different habitat types, and also due to non-randomly missing seine hauls)(references: Clarke 1993; Walters and Coen 2006)

a. Run ANOSIM on [transformed] Bray-Curtis similarity matrix.
b. Need to transform data so super abundant taxa (Anchoa, Bairdiella) don't dominate the results. Probably **square root [**~~or fourth-root transformation (these are common ones; find citation). How to decide which?~~]
    i. Clarke and Green (1988) have a good discussion about possible transformations: square root transformation seems like our best bet at walking the line between not ignoring medium-to-less common species, but not overly weighting them.
    ii. After testing a few transformations, square-root does seem to walk the line best between reducing the orders-of-magnitude spread in data and not being so severe as to reduce to essentially a presence-absence matrix (see McCune and Grace 2002 for more on the severity of anything beyond square-root transformations, and the file '07_transformations_revisited.html' for my graphs)
c. Need to remove rare species? *Updated thoughts: no need to remove; McCune and Grace even make the argument that removal is bad in hypothesis testing.* ~~If so, what's defined as 'rare'? (believe some refs have removed any taxa that didn't make up at least 1% of total catch – abundance? – over the project. Need to find these citations though.)~~
    i. Bray-Curtis similarity is based on species in common between two sites (e.g. Gotelli and Ellison Table 12.6; Wikipedia article is also very good at explaining calculations) – so this shouldn't matter too much for our intended procedures. It does affect the calculation a little bit but shouldn't be huge (only in denominator; numerator is only species in common)
    ii. Transformation seems to matter more than removal for rare vs. common species and abundances (e.g. Clarke and Green 1988; Poos and Jackson 2012)

Figure 1. Example of statistical examination section with more thorough thoughts, justifications, and updates than would belong in a publication or a scope of work document.

Stat
The ISI's Journal for the Rapid Dissemination of Statistics Research
(wileyonlinelibrary.com) DOI: 10.100X/stat.???

Sample ID	Date	Site	Habitat Type	*Anchoa mitchilli* (bay anchovy)	*Lagodon rhomboides* (pinfish)	*Menidia beryllina* (inland silverside)
NFM07-117	2007-07-26	5	beach	0	0	268
NFM07-118	2007-07-26	6	erosional edge	98	0	5
NFM07-120	2007-07-26	8	seagrass	2	30	8

Figure 2. Example spreadsheet of fish monitoring study data with columns for 'Sample ID', 'Date' of sample collection, 'Site' and 'Habitat Type' sampled, and columns of 3 example species (out of nearly 100) with the number of individuals caught in each sample: *Anchoa mitchilli* (bay anchovy), *Lagodon rhomboides* (pinfish), and *Menidia beryllina* (inland silverside).

Season	Habitat	ANOSIM R	ANOSIM p-value	NMDS Stress	SIMPER primary species
Spring	**Overall (global)**				n/a
	beach-depositional			n/a	
	beach-erosional			n/a	
	beach-shell			n/a	
	…			n/a	
Summer	**Overall (global)**				n/a
	beach-depositional			n/a	
	beach-erosional			n/a	
	beach-shell			n/a	
	…			n/a	

Figure 3. Example table for seasonal statistical results: from ANOSIM, dissimilarity measure R and p-value; NMDS Stress measure; and primary species identified by SIMPER as contributing to pairwise dissimilarities. Blank cells will be filled in with values from statistical output; n/a represents output that will not be produced.