

Data Analysis (needs to go in the methods section of your proposal for the second submission)

How will the collected data be analyzed to answer your research question? Where appropriate, which statistical tests will be used to evaluate your hypotheses? Justify your choices by comparing to previously published research.

What needs to go into the Data Analysis section of your proposal? Write at least one sentence on each of the following areas in your Data Analysis:

1. How are you going to express your data and units (example: mean height of students will be expressed in centimeter)... others: %, totals numbers, etc.
 - Results for “x” will be expressed as mean +/- SEM, etc
2. What regression/correlation method(s) will be used to compare the strengths and relationships between your dependent/independent variables. Give ranges for correlation, etc...
3. What statistical test will be used to assess the significance of the data (t-test, Chi Sq, ANOVA, etc).
4. What is your p-value for significance ($P < 0.05$, $p < 0.001$, etc)
 - The null hypothesis will be rejected if $p < 0.05$.

1

Common statistical tests

1. Chi square Test

Determines if certain dependent categorical characteristics (behavior) in two or more samples are different from each other. Example: What is the relationship between gender (M or F) and college majors (Chem, Bio, Math, Environ, etc)?

2. Student' t-Test

Test whether the averages (means) of a certain measured variable of two groups are different from each. Example: do cows given growth hormone have a higher rate of milk production than cows fed with no hormone?

3. **ANOVA (or Analysis of Variance)** is a test applied to compare samples among more than two and up to any manageable number of groups Example: the best concentration of hormone (none, low, medium, high) for milk production in cows?

4. Regression and Correlation

- a. Correlation coefficient -- how strong two variable correlate with one another ranges from -1 to 0 or +1 to 0 BUT DOES NOT TELL you the exact relationship between variables. (see board)
- b. Linear regression – tells you the exact mathematical relationship between two variables, but Not strength (see board)

2

Choosing the right statistical test

- **P-values:** The probability that an event is random OR the probability that the Null hypothesis CANNOT be rejected. *(see week #9 notes on p-value!)*
- **Statistical tests** determine if an event is random/significant by determining the p-value.
- The **independent variable** is typically the variable being manipulated or changed and the **dependent variable** is the observed result of the independent variable being manipulated.
- Choice of statistical test depends on 1) whether **your dependant** variable is **categorical** (ex, classification) **continuous** (measured number), or **ordinal** (ex, rank, score) and 2) the number of groups you have.
 - With **Categorical** variables we usually record the frequency (mode) of various classifications for the independent variables. Example: gender of students in this class versus college majors.
 - With **Continuous** variables, there is equality of intervals in the measurement (a measured number) and we usually determine the “mean”. Example: the height, age, and cholesterol levels of male/female students at FGCU. (use parametric statistical tests)
 - With **Ordinal** variables there is some sort of order, or rank, but no intervals. Somewhat in between Categorical and continuous....such as income levels in terms of low, medium, and high income of male and female student. (use nonparametric statistical tests)
- Determine if your dependent variable(s) are categorical , ordinal, or continuous
- Determine how many groups you have : one, two, three, etc

3

Choosing the right statistical test

Dependent Variable Type

# of GROUPS	Categorical	Ordinal (nonparametric)	Continuous (parametric)
ONE	Binomial test or CHI Square Test	Wilcoxon sign-rank test	t -test
TWO	CHI Square Test	Mann-Whitney U test	t -test
THREE or MORE	CHI Square Test	Kruskal-Wallis test	Analysis of variance (ANOVA)
Correlation /regression Method	Contingency table (contingency coefficient) Simple logistic regression	Spearman correlation Nonparametric regression	Pearson Correlation Coefficient Linear Regression

***Determine the statistical test(s) and correlation method(s) for your proposals.

*** Write your Data Analysis section (include unit of expression, correlation, statistical tests, and p-value --- see earlier slides)

***Follow examples of relevant research papers that you used for justification of your methods and/or introduction.

4

SELECTING YOUR STATISTICAL TEST	Type of Data for your dependent variable		
GOAL	Measurement (from Gaussian Population) CONTINUOUS measured variable (Parametric tests)	Rank, Score, or Measurement (from Non- Gaussian Population) Ordinal (Nonparametric tests)	Binomial (Two Possible Outcomes) Categorical
Describe one group	Mean, SD	Median, interquartile range	Proportion
Compare one group to a hypothetical value	One-sample t test	Wilcoxon test	Chi-square or Binomial test **
Compare two unpaired groups	Unpaired t test	Mann-Whitney test	Fisher's test (chi-square for large samples)
Compare two paired groups	Paired t test	Wilcoxon test	McNemar's test
Compare three or more unmatched groups	One-way ANOVA	Kruskal-Wallis test	Chi-square test
Compare three or more matched groups	Repeated-measures ANOVA	Friedman test	Cochrane Q**
Quantify association between two variables	Pearson correlation	Spearman correlation	Contingency coefficients**
Predict value from another measured variable	Simple linear regression or Nonlinear regression	Nonparametric regression**	Simple logistic regression* 5