Biostatistics Data Memo _____

1) Exploratory Basic Statistics:
   For each group, calculate **the Mean, Median, Variance, SD, Standard Error of Mean(SEM), Standard Error of the Variance(SEV)**, **Confidence Interval, Confidence Bounds** etc. etc.. Do biostatistics basics first.

2) In the event of a bimodal distribution, with large modes, perform correlation testing.
   a. First, do a **unimodality test** to confirm the result. A detailed analysis must be performed by __ (___and ___)in this event, which is not entirely unlikely. Is it due to EF parameter changes, different mortality rates among groups, etc. etc.
   b. We can begin to analyze it by assuming it's the sum of 2 normal distributions. The goal is to estimate the 2 means, the 2 variances (and SD's by extension) and the mixing parameter. Potentially Skew and Kurtosis as well if we want to be stringent.
      Unfortunately, most of the methods of estimating the parameters are numerical in nature. **MLE, Method of Moments**, **EM algorithm** etc.
      This probably best done using R and the mixtools package which are relatively well documented and can execute the EM algorithm with minimal difficulty. Both of which are free.

      Ultimately, these identified parameters are just a tool for us to better understand the underlying pathology of infection in ___and the test system which is a job for the researchers and investigators at ___).

3) Hypothesis testing for a group with 2 samples and a continuous outcome
   a. Normal
      i. Confirm normality via **Anderson Darling and/or Shapiro Wilk** test. For example, the Weibull distribution which looks roughly normal, "gives a distribution for which the failure rate is proportional to a power of time" – a commonly used free online encyclopedia
      Let's make sure we're not modelling filter failure rates for instance. We can calculate the AD statistic against a 3 parameter Weibull distribution. This might be hard to disprove because of the closeness of the Weibull distribution to the normal distribution.
      ii. Do a **T Test.** Calculate the T-value on an Excel sheet, look up the critical value for p=0.05 and or smaller values. Use that to reject or accept the Null hypothesis, namely that _____.
      iii. Do a **T Tests** or an **ANOVA**. Which type is most suitable has to be evaluated.
      One-way Anova, **One-way repeated measures ANOVA**, Factorial Anova, Mixed Anova, MANOVA, ANCOVCA. They only have slight differences but most will not sense based on the planned experimental design and the number of proposed groups. Factorial ANOVA for instance.
   b. Not Normal
      i. (optional but highly recommended) attempt to fit a well known distribution via goodness of fit tests, **Kolmogorov Smirnov, MLEs**, **Anderson Darling statistics**, **P values** etc.
      Come up with a reasonable lists of distributions which we could be looking at. Everything from log normal to exponential, gamma, beta etc. Note that gamma is actually quite likely for this particularly experiment. See _____ for more information.
      ii. (optional but highly recommended) Once a distribution is determined, research what the ramifications of that distribution are. For example, normality/poisson-ness imply a degree of randomness but a Gamma distribution might imply something about the mortality and survival rate of bacteria.
      (https://www.ncbi.nlm.nih.gov/pubmed/28800819)
      iii. Do a **Wilcoxon Matched Pair test**. Calculate the W value, compare it to the critical value and accept/reject the null hypothesis based on that for p=0.05 or smaller. This works surprisingly well because what we're doing is analogous to a matched case study. Namely that we're matching groups to controls, and just comparing against them and it makes no

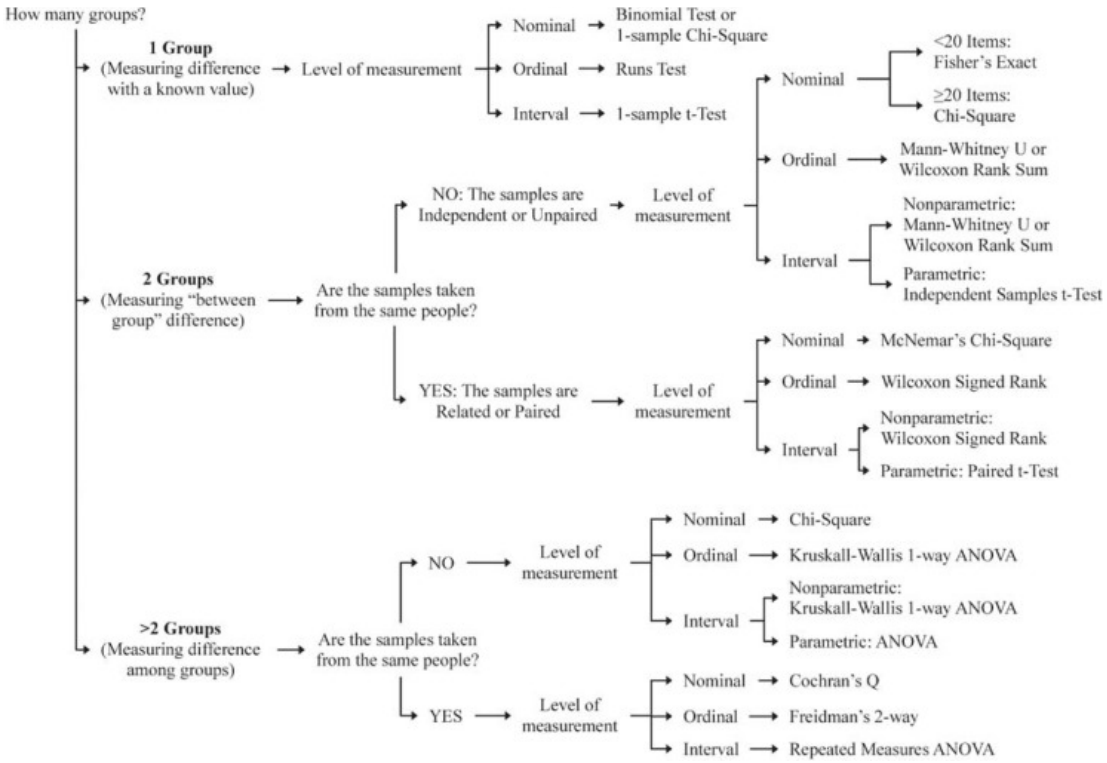
assumptions about the shape of the underlying distributions. Definitely report the result from this test.

4) Correlation Testing – (The technical term for what we're trying to prevent is Confounding) we may not need to publish this but we will need something more conventional than _____ to prove that there is no correlation between certain key environmental factors/variables and plate counts. There are three commonly available options
    a. **Pearsons R**
    b. **Kendall's Tau**
    c. **Spearman's Rho**

    Depending on what the two parameters in question are, based on pathology, it may make sense to use different statistics to test for correlation.
    Pearson's R measure the linear correlation between two variables. Spearman's Rho is in some ways a generalization of that. Instead of looking for a linear correlation, it models how well the relationship between the two variables can be described by a monotonic function. Similar to Spearman's Rho, Kendall's Tau also attempts to measure the rank correlation but not via a monotonic function.

    Depending on what rank correlation you want to prove or disprove, you may want to use one or more of these coefficients to demonstrate correlation or the lack thereof. Consider F Tests and Mutual Information Regression Statistics as well.

5) Calculate Power and Effect Size
    After the T Test and/or Wilcoxon Matched Pair test in part 3, assuming the null hypothesis is significant, you know have to show the result had a large effect. A small effect is more likely to be attributed to sampling differences so a large effect size is always desireable. There are many measure of this: **Glass's Delta, Cohen's D and Hedge's G** which are all based on variation between the means of the groups. (These were the effect sizes I used to calculate the axial sample size). Considering the mean reduction seems reasonable in our case, however, this does not always make sense. The differences between the measures are how they account for variance.
    In the event that we end up finding that the variances between the initial population in each group are different for some reason(which they really shouldn't be because if they are, we've made a protocol error during setup. You can verify this with a **Mann Whitney U test** but that's a bit much), one possible statistic we can use is **Z/sqrt(n)** where Z is the Wilcoxon test statistic. Before using that though, it's probably wise to read :
    Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis. (pp. 231-244). New York: Russell Sage Foundation.
    so that the effect size can be readily interpreted for "usefulness". Note that z/sqrt(n) unlike Cohen's D does not necessarily indicate the magnitude phenomenon in question. It's a statement about the probability that the two groups are in fact different.
    For more information, see:
    https://stats.stackexchange.com/questions/342987/effect-size-for-wilcoxon-signed-rank-test-that-incorporates-the-possible-range-o



https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4552232/

You might want to read this.