QUESTION 1

1. A p-value is:

   The probability that you would observe a statistic equal to or more extreme than yours assuming that the data is random.

   The probability that your alternative hypothesis is true.

   The probability that you would have detected a significant difference if it was present.

   The probability that you would observe a statistic equal to or more extreme than yours assuming that the null hypothesis is true.

QUESTION 2

1. Select the explanation that is NOT a cause of the reproducibility crisis.

   Pressure on academics to publish papers

   statistical software like R

   The file drawer problem

   Small sample sizes

QUESTION 3

1. You grow plants in soil collected from three different habitats the variables you record are soil type and leaf area. Select the description that best describes what type of variables these are. (soil type, leaf area)

   ordinal, discrete

   discrete, nominal

   discrete, continuous

   continuous, discrete

QUESTION 4

1. The mean height of all students that attend Texas A&M University is an example of a _____ . If we go and measure 100 students and calculate the mean height of this group this is a _____. We represent these two values with special letters _____, _____ respectively.

   parameter, statistic , $\bar{Y}$, μ

   statistic, parameter, μ , $\bar{Y}$

   parameter, statistic, μ , $\bar{Y}$

   statistic, parameter, $\bar{Y}$, μ

QUESTION 5

1. If you have the time or funds to increase the number of replicates you can add either biological or technical replicates. _____ replicates are usually most powerful however _____ replicates are particularly useful if your measurements are particularly prone to noisy or extreme measurements.

   technical, biological

   biological, technical

QUESTION 6

1. In an observational study the experimenter assigns individuals into different treatment groups to understand the impact of a possible predictor variable on some response variable of interest.

   True

   False

QUESTION 7

1. We should think deeply about how we sample. Our sampling design (e.g. the strains of an organism that we use, the field sites that we visit, etc.) determines how broadly what we learn during our statistical analysis can be applied.

   True

   False

QUESTION 8

1. You should try several transformations on each variable to see if any of them improve the p-value that you get when running a statistical analysis.

   True

   False

QUESTION 9

1. Which is NOT a benefit of learning and using R.

   It often has the latest most powerful approaches to data analysis that may not be available elsewhere.

   Allows you to have a record of all the analyses that you have done.

   It will allow you to produce high-quality plots for publications and presentations.

There is no learning curve and you don't really need to practice.

QUESTION 10

1. Which is NOT a common problem in plots

   Plots that use axes ranges that exaggerates the difference in measurements among groups.

   Plots that use many colors that don't provide any additional data for your reader.

   Plots that hide the distribution of the data and only show the mean value of many measurements.

   Plots showing both the data as well as some summary like the credible interval of a statistic.

You are interested in whether different majors have different probabilities of accepting a covid-19 vaccine to do this you survey 200 students from each of 5 majors.
For questions 11 and 12 use the dataset vac.rates.csv

QUESTION 11

1. What test did you use to analyze the vaccine acceptance rates among majors data?

   binomial test

   linear model

   anova

   t-test

   chi-square test

   correlation test

QUESTION 12

1. What p-value was associated with your test of the vaccination rates among majors?

You sample birds from two field sites. At each site you count the number of nasal mites that are found in each bird. Determine whether the parasite load is the same for your two samples of birds.
For questions 13 and 14 use the dataset mite.count.csv

QUESTION 13

1. What test did you use to analyze the mite count data?

   linear model

   chi-square test

   t-test

   binomial test

   anova

   correlation test

QUESTION 14

1. What p-value was associated with your test of the mite count data?

You are interested in the impact of different feeds on the growth of chickens. Identify the feed that leads to the highest mean weight and all other feeds that are not significantly different from this feed.
For questions 15 and 16 use the dataset chicks.csv

QUESTION 15

1. Which feed leads to the highest mean chick weight?

   casein

   horsebean

   sunflower

   meatmeal

QUESTION 16

1. Select the feed that led to the highest weight as well as any other feed(s) that are not significantly different from it.

   meatmeal

   horsebean

   casein

   sunflower

   soybean

linseed

You are interested in what determines the reproductive success of betta fish. You have recorded size, color intensity, and strain information for each fish. Determine which variable(s) is predictive of reproductive success.
For questions 17 and 20 use the dataset betta.csv

QUESTION 17

1. What test did you use to analyze the betta fish reproduction dataset?

    anova

    chi-square test

    t-test

    binomial test

    correlation test

    linear model

QUESTION 18

1. Choose the answer that is most correct with regard to the impact of fish strain on reproductive success.

    The Spade strain produces significantly more offspring than any other strain.

    The crown strain produces the fewest offspring.

    The strain has no impact on the number of offspring produced.

    The Plakat and Veil strains have fewer offspring.

QUESTION 19

1. Choose the best description for the impact of body size on reproductive success.

    Body size is not a significant predictor of reproductive success.

    For each unit of increase in the size of the fish, an increase of 0.28 is expected in the number of offspring.

    For each unit of increase in the size of the fish, an increase of 2.55 is expected in the number of offspring.

For each unit of increase in the size of the fish, an increase of 0.05 is expected in the number of offspring.

QUESTION 20

1. Choose the best description for the impact of color intensity on reproductive success.

   For each unit of increase in color intensity, an increase of 16 is expected in the number of offspring.

   For each unit of increase in the size of the fish, an increase of 0.25 is expected in the number of offspring.

   Color intensity is not a significant predictor of reproductive success.

   For each unit of increase in color intensity, an increase of 0.05 is expected in the number of offspring.

EXTRA CREDIT - QUESTION 21

1. Based on your model how many offspring would you expect to get if you were to breed a fish with the following characteristics?

   Body length = 8

   Color intensity = 110

   Strain = crown

   Give your answer as the nearest whole number.