

Download the “NBA2022-2023 regSeason” dataset. This dataset includes per game team statistics for all 30 teams in the NBA for the 2022-2023 regular season and whether they made the playoffs.

Perform both one-sided and two-sided t-test analyses (assuming variances equal) to determine which NBA stat categories were significantly different for teams that made the playoffs vs. those teams that did not make the playoffs. Report your results in Table form. Report the R^2 , N, p-value, significance (Y/N). The table is started for you:

NBA stat	Two-sided test p-value	One-sided test p-value (Prob > t)	Significant difference in Means (Y/N)	Significant increase in mean (Y/N)	% variation explained
MP	0.3041	0.1520	N	Y	3%
FG	0.0904	0.0452	y	Y	9%

Which NBA stat categories were significant ($\alpha = 0.05$), and which were close to significant ($0.10 > p > 0.05$) for the one-sided t-test? What does this mean?

there is a significant difference between the variables of the one-side t-test

Which NBA stat categories were significant ($\alpha = 0.05$), and which were close to significant ($0.10 > p > 0.05$) for the two-sided t-test? What does this mean?

Repeat this analysis (and provide another table like above) but conduct your t-tests assuming variances are unequal. Are there any major differences regarding which NBA stat categories are significant? when the variances are unequal, the results of the t-test may differ slightly from the results of the standard two-sample t-test assuming equal variances.

Go back to the overall JMP datafile for this dataset. Rank the NBA stat categories that **were significant from above (assuming unequal variances)**. You can do this by using the Distribution function for a given column->selecting the lower red arrow for a given distribution->save-> ranks. This will produce a ranking column in your dataset for that NBA stat category with the higher the number in the distribution representing a higher rank (e.g., the team with the highest FG% will have a rank score of 30). Once you have made the rank columns, create a new formula column in which you sum all the ranks from each column and divide by the total to get the overall average rank. You can call this column the Overall Rank.

Since you found these stat categories in the regular season to be significant, use the Overall Rank that you just calculated to predict which team may have the best chance of winning the 2023 NBA Finals based on these regular season stats. Which teams are predicted to have the second and third best chance of winning the Finals based on this analysis of Overall Rank? Give that this analysis is based off of regular season stats, why may these predictions not actually be that accurate for the postseason (playoffs)?

Provide your data table that includes the rank columns and your Overall Rank column.

Go back to the quillback images from earlier in the class. Measure the eye diameter of each fish and add this to your data table. Calculate the eye diameter to head length ratio. Go to the canvas section where the quillback images are found and find the bigmouth buffalo images, as well as the redhorse images. Download these images and measure both eye length and the head length as you did for quillback, and then use ANOVA to determine if the ratios significantly differ by species. Then, measure the total body area of each bigmouth buffalo and redhorse as you did for quillback in an earlier assignment, and count the lateral line scales. Use ANOVA once again to determine if the body areas between the three species significantly differ, and whether the number of lateral line

12

scales significantly differ. Use Tukey honest significant differences (HSD) post-hoc test like we have done before and use the Connecting Letters Report.

Show your JMP stat outputs and your data table. Interpret the ANOVA results in context by providing concluding statements about each test that you performed.

Find fishID248 from the image files, a female bigmouth buffalo, and measure its total length. Using what you know from Figures 6 and 7 from [this study](#), what would you predict for this bigmouth buffalo's age? Why?

Extra Credit (12 pts):

Redhorse (*Moxostoma* spp.) are a diverse genus of native river fish found in North America. Several are native to Minnesota. Download the "MNredhorse..." dataset. Answer the following questions according to this dataset:

How many species of redhorse are found in MN? (1/4 pt)

How many redhorse individuals have been measured for LEN_MM (total length) in this dataset? (1/4 pt)

What years does this dataset span? (1/4 pt)

Which redhorse species is most abundant in Minnesota? (1/4 pt)

Describe the distribution shape of LEN_MM for redhorse overall (1pt).

Use your knowledge from HW10 of the bubble plot graph option, make a map figure of the distribution of redhorse by species across years, using size in mm as the color gradient. Hint: you will also need to create an ID column so that each row in the dataset has a unique ID. Make sure US State boundaries are shown, and you can use whichever background map type (street, simple, detailed, etc.) that you prefer. Upload gifs for each species map across time as separate upload file attachments on Canvas (Mac Users will need to use the screen capture command (command+shift+5) and proceed to record the playing of the animation). Mac users do not need to upload the captures as separate files – Mac users can insert them as screen captures into the Word doc and they should still be able to play, however check this to make that it works. Keep each Gif or screen capture recording under 10 seconds. You do not need to do one for Black Redhorse because there is not enough data for that species. (7 pts)

Research a little about shorthead redhorse and greater redhorse size. Go back to the dataset and investigate shorthead redhorse. What do you notice for shorthead redhorse that must be an error in this dataset? What must be an error for one particular Golden Redhorse? (3pts)